

Ghent University

Faculty of Science

Department of Molecular Genetics

MOLECULAR CHARACTERIZATION OF TEMPERATURE-SENSITIVE GROWTH MUTANTS OF *ARABIDOPSIS THALIANA*

Elke Vaeck

Thesis submitted in partial fulfillment of the requirements for the degree of

Doctor (Ph. D.) in Science (Biotechnology)

Academic year 2005-2006

Promotor: Prof. Dr. Marc Zabeau

Co-promotor: Prof. Dr. Dirk Inzé



EXAM COMMISSION

Promotor

Prof. Dr. Marc Zabeau (Secretary)
Dept. of Plant Systems Biology
Flanders Interuniversity Institute for Biotechnology (VIB) – Ghent University

Co-Promotor

Prof. Dr. Dirk Inzé
Dept. of Plant Systems Biology
Flanders Interuniversity Institute for Biotechnology (VIB) – Ghent University

Promotion Commission

Prof. Dr. Rudi Beyaert (Chairman)
Dept. of Molecular Biomedical Research
Flanders Interuniversity Institute for Biotechnology (VIB) – Ghent University

Dr. Marnik Vuylsteke
Dept. of Plant Systems Biology
Flanders Interuniversity Institute for Biotechnology (VIB) – Ghent University

Dr. Gerrit T.S. Beemster
Dept. of Plant Systems Biology
Flanders Interuniversity Institute for Biotechnology (VIB) – Ghent University

Dr. Frank Van Breusegem
Dept. of Plant Systems Biology
Flanders Interuniversity Institute for Biotechnology (VIB) – Ghent University

Prof. Dr. Wout Boerjan
Dept. of Plant Systems Biology
Flanders Interuniversity Institute for Biotechnology (VIB) – Ghent University

Dr. Janny L. Peters
Dept. of Plant Genetics
Institute for Wetland and Water Research – Radboud University Nijmegen

Dr. Sandra Ormenèse
Laboratoire de Physiologie Végétale
Université de Liège (Sart-Tilman)

*If the doors of perception were cleansed
everything would appear to man as it is, infinite.*

W. Blake

Acknowledgements

Spending five years preparing my PhD thesis was a challenging but rewarding experience. During this period, my colleagues, family and friends have supported me in many different ways. For this, I want to express my sincere appreciation to them.

First, I want to thank my promotor, Marc Zabeau, for giving me the opportunity to start working in the department and for expressing confidence in me. Marc, during the past five years you were always there at the crucial moments to give useful input, advice and new, inspiring insights.

Second, I'm grateful to Dirk Inzé for allowing me to work in a modern and well-equipped research environment and for his continuous support and advice.

Also, I want to express my gratitude to Pia Neyt, Janny Peters and Tom Gerats for the help and guidance they gave me during the first year of my PhD research. After Janny and Tom left the lab, Gerda Cnops further assisted me in the map-based cloning procedure, which I have sincerely appreciated.

From the second year on, Marnik Vuylsteke guided me in my work and recently helped me a lot in finalizing this thesis. Marnik, thank you for your input and help, the useful conversations, the trust, and ... the statistics.

Gerrit Beemster has taken the responsibility of being my "second advisor" during the past two years, especially when it came down to the biological stuff. He also spent quite some time to work his way through this thesis. Gerrit, thank you!

In Marnik's lab, I met a lot of new people over the years, some of which have become real friends. Hilde, Lindy, Ruth, Ilse, Emmy, Debbie, Raphaël, Roel, Jeroen, Annelies and others, thank you for the enjoyable time together!! Especially, I want to mention Emmy Coppens, who, with her dedicated work, made a substantial contribution to the map-based cloning of the ts mutants.

I also want to thank my husband, Kurt, who has supported me fully all the way from the beginning of this work and who especially during the last six months, had to share with me all the inconvenience and stress of a race to the deadline for finalizing this thesis. I could not imagine succeeding in this work without the support of my wonderful parents, who always encouraged me to go for new challenges in life. My very sincere gratitude to my little sister, Kirsten, and brother in law, Andy, who were always there for me in difficult times and who developed the design for the cover of this thesis.

THANK YOU!

Table of Contents

Chapter One

GENERAL INTRODUCTION

1.1 The Identification of Genes with Unique and Essential Functions	2
1.2 Temperature-Sensitive Mutants: Tools to Study Essential Genes <i>In Vivo</i>	2
1.3 Establishing Gene Function Through Mutagenesis	19
1.4 Plant Functional Genomics Through Transcriptome Analysis	46

Chapter Two

ISOLATION OF TEMPERATURE-SENSITIVE MUTATIONS IN *ARABIDOPSIS THALIANA*

2.1 Introduction	58
2.2 Experimental Procedures	59
2.3 Results	62
2.4 Discussion	70
2.5 Conclusion	74

Chapter Three

MAP-BASED CLONING OF THE TEMPERATURE-SENSITIVE MUTATIONS *dgr1* AND *dgr2* IN *ARABIDOPSIS THALIANA*

3.1 Introduction	77
3.2 Experimental Procedures	78
3.3 Results	82
3.4 Discussion	90
3.5 Conclusions	101

Chapter Four

EXPRESSION ANALYSIS OF THE TEMPERATURE-SENSITIVE *dgr2* MUTANT OF *ARABIDOPSIS THALIANA*

4.1 Introduction	104
4.2 Experimental Procedures	104
4.3 Results	107
4.4 Discussion	112
4.5 Conclusion	114

Chapter Five

GENERAL CONCLUSIONS AND PERSPECTIVES

114

Chapter Six

SAMENVATTING

117

References

120

CHAPTER ONE

GENERAL INTRODUCTION

In little more than a decade, research in plant biology has taken huge steps forward towards understanding nearly all aspects of plant growth and development. Although significant progress has been made in many species, we will focus here on *Arabidopsis thaliana*, which has been adopted as a model system for molecular genetic studies. One of the long-term goals of *Arabidopsis* research is to obtain a better understanding of all the biological processes that operate in a plant. To achieve this, it is necessary to define the minimal set of genes required to produce a viable plant with a normal phenotype under diverse conditions. Because of the complexity and the multitude of biological processes required for a plant to grow and develop, a large and diverse set of essential genes are likely to be involved. Applying forward genetics is a powerful method to identify the relevant genes. This approach usually involves a screen for mutants that are used to identify genes involved in specific processes. However, if the genes underlying these processes are essential for the viability of the plant, mutations will often result in a lethal phenotype. Therefore, conditional lethal mutants are very useful tools to study the function of essential genes *in vivo*.

In our own work, we isolated temperature-sensitive *Arabidopsis* mutants to identify novel genes that are essential for plant growth and development. Temperature-sensitivity is the most common type of conditional mutation in which the mutant phenotype only appears at certain (usually higher) restrictive temperatures. This way, gene function can be studied at the restrictive temperature, while mutant lines can be maintained and propagated at the permissive temperature, at which the wild type phenotype is expressed. Through the phenotypic analysis and molecular genetic characterization of these temperature-sensitive mutants we aimed to unravel a piece of the huge network of biochemical processes that underlies plant growth and development. This introductory chapter presents an overview of the history, use and generation of temperature-sensitive mutants and the molecular genetic methods and technologies that are currently applied for their characterization.

1.1 THE IDENTIFICATION OF GENES WITH ESSENTIAL FUNCTIONS

'Essential' and 'non-essential' are classical molecular genetic designations that relate to the functional significance of a gene with respect to its effect on organismic fitness. Generally, a gene is considered to be essential if a knockout results in (conditional) lethality or infertility. In organisms of which the genome sequence is known, the large-scale identification of essential genes with unique functions using gene inactivation technologies has become an important research topic. The long-term goal is to define the minimal gene set needed to produce a viable organism with a normal phenotype under diverse conditions. The systematic identification of essential genes has been described in several prokaryotes, including *Mycoplasma genitalium* (Hutchison et al., 1999), *Staphylococcus aureus* (Ji et al., 2001), *Haemophilus influenzae* (Akerley et al., 2002), *Vibrio cholerae* (Judson et al., 2000) and *Bacillus subtilis* (Kobayashi et al., 2003). Related work has also been published for two model eukaryotes, *Saccharomyces cerevisiae* (Giaever et al., 2002) and *Caenorhabditis elegans* (Kamath et al., 2003) and significant advances have been reported in the identification of essential genes associated with human diseases (Jimenez-Sanchez et al., 2001).

Only recently, a comparable dataset of essential plant genes has been constructed (Tzafrir et al., 2004). An initial collection of 250 EMB genes required for normal embryo development in *Arabidopsis* is described and represents the first large-scale dataset of essential genes in a flowering plant. Sixty percent of these EMB identities were derived from a forward genetic screen of T-DNA insertion lines generated at Syngenta (McElver et al., 2001). Most EMB genes are required for viability and may therefore be considered as essential in a traditional sense. When compared with 550 genes linked to other knockout phenotypes, EMB genes are

enriched for basal cellular functions, deficient in transcription factors and signalling components, have fewer paralogs and are more likely to have counterparts among essential genes of yeast (*Saccharomyces cerevisiae*) and worm (*Caenorhabditis elegans*). EMB genes also represent a valuable source of plant-specific proteins with unknown functions required for growth and development. Although the large-scale identification of essential genes is possible using regular knock-out mutants, the determination and study of gene function *in vivo* using such mutants is often problematic due to the resulting lethal phenotypes (Despres et al., 2001).

1.2 TEMPERATURE-SENSITIVE MUTANTS: A TOOL TO STUDY ESSENTIAL GENES *IN VIVO*

Conditional lethal mutants have provided the means to genetically study mutations in essential genes. The potential to control the activity of individual genes reversibly and quantitatively has opened up new perspectives for the in-depth study of gene function. A decisive feature of this system is that phenotypically the state of a null-mutation can be induced without altering the target gene itself, while the possibility to subsequently return to the wild type (WT) state makes the mutant phenotype truly conditional. Equally important is the quantitative nature: gene activities can be varied by small increments, a feature that is of particular interest for the study of gene products that are involved in a variety of intracellular equilibria. Whereas in general they will participate in such interactions with different affinities, a gradual change of their intracellular concentration may reveal phenotypes and targets of interaction that remain obscure when the respective gene is simply inactivated.

Temperature-sensitive (ts) mutations are the most widely used type of conditional mutations in which there is a

marked drop in the level or activity of the gene product when the gene is expressed above (or below) a certain temperature (restrictive temperature) (Chakshumathi et al., 2004). Below (or above) this temperature, at so-called permissive temperatures, the phenotype of the mutant is similar to that of the WT. Ts mutants provide an extremely powerful tool for studying gene function *in vivo*. They represent a reversible mechanism to lower the level of a specific gene product at any stage in the growth of the organism simply by changing the temperature of growth. Although there are a number of other excellent inducible systems for gene expression (Gossen and Bujard, 1992; Gossen et al., 1995) ts mutants have several unique advantages over such systems, including fast temporal response, high reversibility and the applicability to any tissue type or developmental stage of an organism. The use of ts alleles is therefore a well-established method that is applicable to any organism upon which temperature changes can be imposed. These include viruses, prokaryotes, yeasts, nematodes, insects, plants and even amphibia and fish (Zeidler et al., 2004).

1.2.1 IN THE EARLY 50's, THE FIRST TEMPERATURE-SENSITIVE MUTANTS WERE EXPLOITED TO OBTAIN MUTATIONS IN ESSENTIAL GENES

In 1946 Mitchell and Houlahan produced the first published description of a ts mutant (*Neurospora*). Subsequently, Horowitz and Leupold isolated several ts mutations in *Neurospora* (1951). The use of these conditional mutants contributed significantly to the establishment of the one gene-one enzyme paradigm stated by Beadle and Tatum in 1945. Further isolation of ts mutations was performed to estimate the fraction of essential genes in *Escherichia coli*. This work was soon recognized by Edgar and co-workers who demonstrated in bacteriophage T4 that ts

mutations could be used effectively to gain information about essential gene functions and that they can occur in many different genes in an organism (Edgar and Lielausis, 1964; Epstein et al., 1963). From that moment, extensive ts lethal mutant collections were created in bacteriophage T4 (Wood and Revel, 1976) and *Escherichia coli* (Kohyama et al., 1966; Sevastopoulos et al., 1977). The mutations in these collections have allowed the genetic dissection of essential functions such as DNA replication, the cell cycle and cell division, membrane biogenesis, transcription and translation (Allen et al., 1970; Schmidt et al., 1989).

1.2.2 TEMPERATURE-SENSITIVE YEAST MUTANTS HAVE SERVED AS POWERFUL TOOLS FOR THE IDENTIFICATION OF NOVEL CELL CYCLE GENES

Since the discovery of yeast cell division cycle mutants by Hartwell and colleagues in the early 70's, the field of cell cycle research has largely benefited from the use of ts mutations (Cayrol et al., 1997). When Hartwell launched the study of cell cycle genetics nothing was known about the mechanisms for achieving the orderly progression of cell cycle events. He found evidence for genes, *cdc* genes, whose function is required at specific time points in the cycle by identifying mutants that arrested the cell cycle at a specific stage (figure 1.1).

The organism that Hartwell chose to study was the budding yeast, *Saccharomyces cerevisiae*. Because the cell cycle is an essential process, mutagenized yeast cultures were screened for ts mutants that grew and divided normally at the permissive temperature, but arrested cell division with a characteristic phenotype at the restrictive temperature (figure 1.2). The first paper to describe *cdc* mutants was by Hartwell, Culotti and Reid (1970), who also

coined the term 'execution point' - the stage in the cell cycle when the gene function is required. In this paper, three *cdc* genes were described, which paved the way for the identification of many more such genes, and for the discovery of the molecules and mechanisms controlling the cell cycle.

Although routine methods for cloning yeast genes were still ten years away, Hartwell's work went much further. With thoughtful and elegant genetic analysis, he used the mutants as tools to block the cell cycle at specific stages and to ask questions about the interdependence of yeast cell-cycle events. For example, in his *J. Mol. Biol. Paper* (1971), Hartwell concluded that DNA synthesis is required for later events, such as nuclear division. By contrast, the formation of the bud is not dependent on DNA replication, although DNA replication does prevent further bud initiation in the same cell cycle.

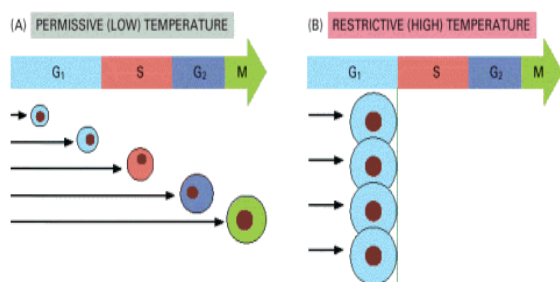


FIGURE 1.1 THE BEHAVIOR OF A *ts* CDC MUTANT (GRIFFITHS ET AL., 1996). (A) At the permissive (low) temperature, the cells divide normally and are found in all phases of the cycle (B) On warming to the restrictive (high) temperature, at which the mutant gene product functions abnormally, the mutant cells continue to progress through the cycle until they come to the specific step that they are unable to complete (initiation of S phase, in this example). Because the *cdc* mutants still continue to grow, they become abnormally large. By contrast, non-*cdc* mutants, if deficient in a process that is necessary throughout the cycle for biosynthesis and growth (such as ATP production), halt haphazardly at any stage of the cycle depending on when their biochemical reserves run out.

Three years further and yeast cell-cycle genetics was well established, as demonstrated by the synthetic model of the cell cycle that was presented in *Science* (1974). Genetics had divided the cycle into two parallel pathways, which comprised two sets of dependent steps and involved a total of 19 *cdc* genes. At the beginning of the cell cycle, both pathways depend on the completion of a step that Hartwell and co-workers termed 'start'. This event is defined by the famous *cdc28* mutant (famous because the *cdc28* gene was later shown to encode the founding member of the cyclin-dependent kinase family), and is also the event at which yeast mating factor arrests cell division to prepare cells for mating. 'Start' turned out to be an important control point for the cell cycle in many eukaryotes.

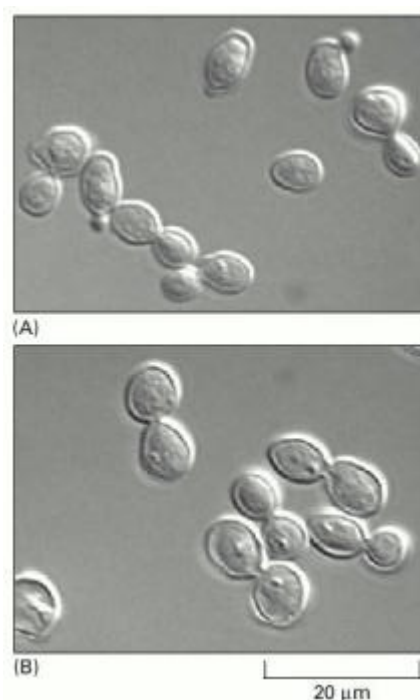


FIGURE 1.2 THE MORPHOLOGY OF BUDDING YEAST CELLS ARRESTED BY A CDC MUTATION (GRIFFITHS ET AL., 1996). (A) In a normal population of proliferating yeast cells, buds vary in size according to the cell-cycle stage. (B) In a *cdc15* mutant, grown at the restrictive temperature, cells complete anaphase but cannot complete the exit from mitosis and cytokinesis. As a result, they arrest uniformly with the large buds, which are characteristic of late M phase.

Later on, ts mutations have also proved invaluable for the dissection of other diverse processes in microorganisms such as *Chlamydomonas reinhardtii* flagellar assembly (Adams et al., 1982; Huang et al., 1977) and protein folding (Goldenberg et al., 1983; Gordon and King, 1993; Gordon and King, 1994; Smith et al., 1980).

1.2.3 VARIOUS ASPECTS OF THE DEVELOPMENT OF *CAENORHABDITIS ELEGANS* HAVE BEEN STUDIED USING TEMPERATURE-SENSITIVE MUTANTS

A diversity of developmental processes has been dissected in the nematode *Caenorhabditis elegans*, in which ts mutations have played a crucial role. Most of these studies describe ts mutations causing cell division and proliferation deficiencies.

Several large-scale screens for embryonic lethal mutants have identified mutations that affect various aspects of cell division in the early embryo (Hirsh and Vanderslice, 1976; Miwa et al., 1980; Cassada et al., 1981; Kemphues et al., 1988a, Kemphues et al., 1988b), while another screen has identified mutations that specifically affect the postembryonic lineages (Horvitz and Sulston, 1980). Both sets include mutations that disrupt the temporal and spatial patterns of cell division as well as mutations that block division altogether. As not all lineages are affected by these mutations, it is not known whether these genes encode key players in the mechanisms of division that are common to all cells.

More recently, O'Connell, Leys and White (1998) designed a genetic screen for ts cell-division mutants to identify nonlineage-specific cell-division mutations in *Caenorhabditis elegans*. They identified genes that encode key components of the cell-division machinery, and that are thus required for most cell divisions, and studied the effects of these mutations during the early embryonic divisions. They

observed defects in chromosome segregation, spindle morphogenesis and alignment, centrosome duplication and cytokinesis. These results confirm that this approach is capable of identifying genes with a broad range of essential functions including roles in cell division.

1.2.4 STUDIES WITH CONDITIONAL LETHALS IN *DROSOPHILA MELANOGASTER*, REVEALED THAT TEMPERATURE-SENSITIVITY IS NOT A PROPERTY CONFINED TO MUTATIONS ON CERTAIN CHROMOSOMES

Ts mutations have been exploited to study a variety of genetic, developmental and neurological phenomena in the model insect *Drosophila melanogaster* (reviewed by Suzuki 1970; Suzuki, Kaufman and Falk, 1976; Hall, 1978a; Hall and Greenspan, 1979) (figure 1.3).

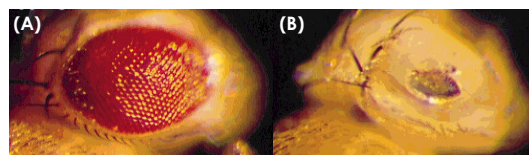


FIGURE 1.3 TEMPERATURE-DEPENDENT DISTURBED EYE PHENOTYPE IN *DROSOPHILA MELANOGASTER* (GRIFFITHS ET AL., 1996). Severe defects are observed when fly mutants are transferred from 22°C (A) to 30°C (B).

During these studies, it seemed that temperature-sensitivity is not a property confined to mutations on certain chromosomes or to a limited number of sites within a chromosome. Indeed, recessive and dominant ts mutations on the X-chromosome (Suzuki et al., 1967; Wright, 1968), Y-chromosomes (Ayles et al., 1972), chromosome 2 (Baillie, Suzuki and Tarasoff, 1968; Suzuki and Procunier, 1969) and chromosome 3 have been analyzed (Fattig and Rickoll, 1972; Holden and Suzuki, 1968). In addition, cold-sensitive recessive lethals on the X-chromosome (Mayoh and Suzuki 1972) and dominant

lethals on the autosomes were isolated (Holden and Suzuki, 1973; Rosenbluth, Ezell and Suzuki, 1972; Suzuki and Procunier, 1969; Tasaka and Suzuki, 1972).

Recently, Barry Ganetzky and Robert Kreber created a collection of more than 100 different *ts* neurological *Drosophila* mutants, which are behaviorally normal at 25°C, but display severe locomotor defects, including uncoordinated movement, ataxia, seizures or complete paralysis, within five minutes of exposure to 38°C. These mutants are being used in screening programs aimed at identifying novel pharmaceutical agents and drug targets in an *in vivo* biological system. They are also used in functional genomics research to identify novel genes involved in neural function and diseases and as experimental models for human neuromuscular disorders, including epilepsies, muscular dystrophies, periodic paralyses, myotonias, dystonias and others (Palladino, Hadley and Ganetzky, 2002).

1.2.5 PLANT TEMPERATURE-SENSITIVE MUTANTS AS TOOLS TO ANALYZE DEVELOPMENTAL AND GROWTH PROCESSES

1.2.5.1 SOMATIC EMBRYOGENESIS IN *DAUCUS CAROTA* (CARROT) CELLS

Ts mutations have been successfully exploited in research into somatic embryogenesis of cultured cells of carrot (Breton and Sung, 1982; Giuliano et al., 1984; Lo Schiavo et al., 1988, 1990; Schnall et al., 1988). Although cultures of carrot cells provide a unique experimental system for studies of somatic embryogenesis, which is a straightforward example of the plasticity of plant cells, further genetic analysis of mutants is difficult because of the poorly characterized genetic background of carrot.

1.2.5.2 PHOTOSYNTHESIS IN *MELILOTUS ALBA* (SWEETCLOVER)

In sweetclover *ts* mutants were isolated in which the restrictive growth temperature produces a chlorophyll deficient phenotype (Bevins et al., 1993; Markwell et al., 1986; Yang et al., 1990). In the past, several approaches have been used in attempts to define the genetic control and biochemical nature of the photosynthetic apparatus in plants. The most frequently used mutants are defective in chloroplast structure, function, and/or pigment biosynthesis. Mutations resulting in the inability to express chlorophyll A are lethal once seed reserves are depleted because this pigment is absolutely required for photochemical energy conversion. When a severe deficiency of chlorophyll a is not lethal, growth is marginal and often results in poor seed production and difficulty in maintaining mutant lines. In contrast to the situation with chlorophyll a deficient mutants, chlorophyll b functions only in a light-harvesting capacity and many healthy mutants defective in the expression of chlorophyll b have been reported. Nevertheless, the utility of mutants that are temperature-sensitive for chlorophyll expression was corroborated since this study demonstrated that the relief of photosynthetic energy limitation that accompanies temperature-induced greening is directly correlated with progressive changes in carbon isotope fractionation, a sensitive physiological indicator. Therefore, the authors advise other researchers interested in chloroplast development to consider the use of *ts* mutants.

Also in *Arabidopsis thaliana* (Markwell and Osterman, 1992), *Zea mays* (corn), *Gossypium hirsutum* (cotton) and *Glycine max* (soybean) (Alberte et al., 1974) *ts* mutants have been used to study the photosynthetic apparatus.

1.2.5.3 CELL GROWTH ANISOTROPY IN *ARABIDOPSIS THALIANA*

Cortical microtubules and cellulose microfibrils have been widely accepted as primary determinants of growth anisotropy in plants (Wasteneys, 2000; Baskin, 2001). To support anisotropic growth, it is essential that microtubules and cellulose microfibrils are oriented at right angles to the main axis of cell expansion, especially in rapidly elongating cells. The cellulose synthase constraint model proposes that cortical microtubules control the direction of cellulose deposition by guiding the cellulose synthase enzyme complexes resident in the plasma membrane as they add cellulose microfibrils to the inner layer of the cell wall (Giddings and Staehelin, 1991). This hypothesis, however, is incompatible with some experimental evidence (Emons et al., 1992; Preston et al., 1988). The alignment of cortical microtubules and cellulose microfibrils is not always similar, even in diffusely expanding cells (Sugimoto et al., 2000).

A number of *ts Arabidopsis* mutants with cell expansion defects have started to provide insights into the cell wall structures that regulate plant cell expansion (Williamson et al., 2001).

In a recent study Sugimoto et al. (2003) tested the cellulose synthase constraint model using two of these *ts* mutants, *rsw1-1* and *mor1-1* (figure 1.4). The *RADIAL SWELLING1* (*RSW1*) gene encodes a cellulose synthase catalytic subunit AtCesA1 (Sugimoto et al., 2001; Williamson et al., 2001) while the *MICROTUBULE ORGANIZATION1* (*MOR1*) gene encodes a microtubule associated protein (Whittington et al., 2001; Eleftheriou et al., 2005). They analyzed cellulose microfibril patterns in cells that had developed either with microtubules completely de-polymerized with the drug oryzalin, or disorganized in the *mor1-1* mutant at its restrictive temperature. Cellulose microfibrils continued to be

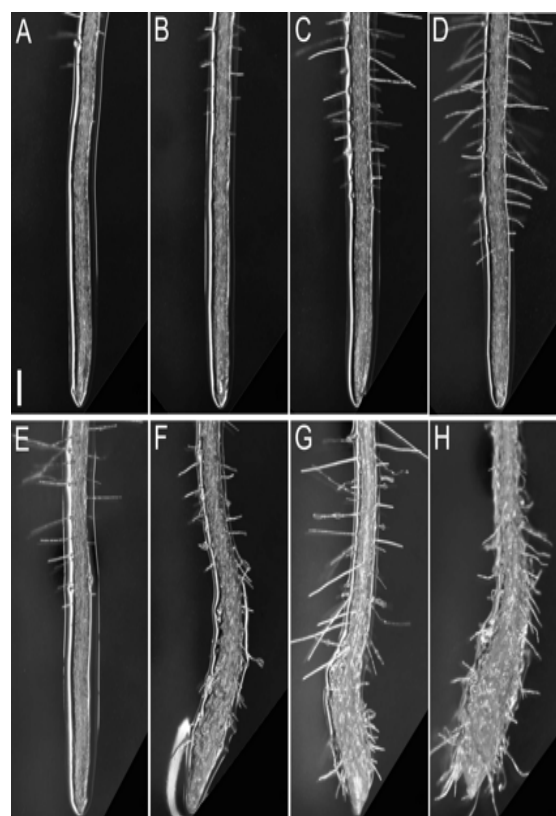


FIGURE 1.4 TEMPERATURE-DEPENDENT MORPHOLOGICAL PHENOTYPES OF *MOR1-1*, *RSW1-1*, AND *MOR1-1-RSW1-1* DOUBLE MUTANT ROOT TIPS (HIMMELSPACH ET AL., 2003). (A) Wild type after 7 days at 21°C, (B) *mor1-1* after 7 days at 21°C, (C) *rsw1-1* after 7 days at 21°C, (D) *mor1-1 rsw1-1* double homozygote after 7 days at 21°C, (E) wild type grown at 21°C for 5 days followed by 2 days at 29°C, (F) *mor1-1* grown at 21°C for 5 days followed by 2 days at 29°C, (G) *rsw1-1* grown at 21°C for 5 days followed by 2 days at 29°C, (H) *mor1-1 rsw1-1* double homozygote grown at 21°C for 5 days followed by 2 days at 29°C. Bar = 100 µm.

deposited in parallel order under these conditions. This result suggested that organized cortical microtubules are not required for oriented deposition of microfibrils and regulate growth anisotropy by some other mechanism. The independence of cortical microtubule organization and cellulose microfibril alignment was supported by the observation that double mutants of *mor1-1* and *rsw1-1*, cellulose-deficient mutants with misaligned microfibrils, had additive phenotypes. Nevertheless, this study did not rule out the possibility that the deposition of microfibrils was simply

maintained because of pre-existing well ordered microfibrils, which, once established, might serve as a template for the continued deposition of transverse, parallel microfibrils (Himmelsbach et al., 2003).

The ts mutants *mor-1*, *rsw1-1* and others like *rsw-2* (Gillmor et al., 2002; Lane et al., 2001; Nicol et al., 1998), -3 (Burn et al., 2002; Peng et al., 2000), -4 and -7 (Wiedemeier et al., 2002) provide a powerful system in which to test the causal relationship between cortical microtubules, cellulose microfibrils and anisotropic cell expansion since it is possible to reversibly alter microtubule organization or cellulose production simply by elevating the temperature by a few degrees.

1.2.5.4 ORGANOGENESIS IN *ARABIDOPSIS THALIANA*

ROOT DEVELOPMENT

To investigate the genetic basis of adventitious organogenesis Konishi and Sugiyama (2003) isolated nine ts mutants defective in various stages of adventitious root formation. Five *root initiation defective* (*rid*) mutants failed to initiate the formation of root primordia. In one *root primordium defective* (*rpd*) mutant, the development of root primordia was arrested. The last three mutants were *root growth defective* (*rgd*) after the establishment of the root apical meristem (Sugiyama, 2003).

The temperature-sensitivity of callus formation revealed further distinctions between the isolated mutants. The *rid1* mutant was specifically defective in the reinitiation of cell proliferation from hypocotyl explants, while *rid2* was also defective in the reinitiation of cell proliferation from root explants. Furthermore, these two mutants exhibited abnormalities in the formation of the root

apical meristem when lateral roots were induced at the restrictive temperature.

The *rgd1* and *rgd2* mutants were deficient in root and callus growth, whereas the *rgd3* mutation specifically affected root growth.

The *rid5* mutant required higher auxin concentrations for rooting at the restrictive temperature, implying a deficiency in auxin signaling. The *rid5* phenotype was found to result from a mutation in the *MOR1/GEM1* gene (§1.2.5.3) encoding a microtubule associated protein, suggesting a possible function of the microtubule system in auxin response.

SHOOT DEVELOPMENT

The isolation of three ts mutants in *Arabidopsis* that are defective in the redifferentiation of shoots from root explants has been described by Yasutani et al. (1994). In an attempt to genetically dissect shoot redifferentiation and at the same time also root redifferentiation the responses in tissue culture of the *srd* mutants were characterized (Ozawa et al., 1998). Shoot and root redifferentiation can be induced at high frequency in *Arabidopsis* from hypocotyls and root explants by a two-step culture method. Tissues are precultured on callus-inducing medium and then transferred onto shoot- (for shoot redifferentiation) or root- (for root redifferentiation) inducing medium. The *srd* mutants exhibited temperature-sensitivity at different steps of organogenesis, which allowed the identification of three states associated with organogenic competence: 'IC' (incompetent), 'CR' (competent with respect to root redifferentiation), 'CSR' (competent with respect to shoot and root redifferentiation) (Sugiyama, 1999).

Hypocotyl explants were shown to be in the 'IC' state at the initiation of culture and to enter the 'CSR' state, via the 'CR' state, during preculture on callus-inducing medium. Root explants seemed

to be in the 'CR' state at the initiation of culture. The transition from 'IC' to 'CR' and that from 'CR' to 'CSR' appeared to require the functions of *SRD2* and *SRD3*, respectively. Moreover, explants in the 'CSR' state redifferentiate shoots with the aid of *SRD1* and *SRD2* when transplanted onto shoot-inducing medium. Histological examination of the *srd* mutants revealed that the function of *SRD2* is required not only for organogenesis but also for the reinitiation of cell proliferation in hypocotyl explants during culture on callus-inducing medium. Linkage analysis using RFLP markers indicated that *SRD1*, *SRD2* and *SRD3* are located at the lower region, the central region and the upper region of chromosome I, respectively.

Making use of the *srd* mutants a hypothetical scheme for *in vitro* organogenesis of *Arabidopsis thaliana* could be set up. The process of organogenesis was subdivided into genetically distinct phases. These phases could be related to physiologically identified phases by tissue-transfer experiments. In addition, such mutants help to find elementary processes that are common to apparently different phenomena.

In another study Pickett et al. (1996) designed a screen for *ts* mutations that cause a conditional arrest of early shoot development in *Arabidopsis* to identify genes involved in meristem function. They describe the characterization of three mutations: *add* (*arrested development*) 1, 2 and 3 (figure 1.5). Analysis of the *add* mutant phenotypes indicates that these mutations are sensitive to temperature shifts at various points in development, suggesting that their activity is required throughout development.

At the restrictive temperature the *add1* and *add2* mutations disrupt apical meristem function as assayed by leaf initiation. Furthermore, *add1* and *add2* exhibit defects in leaf morphogenesis following upshift from permissive to restrictive temperature. This result suggests

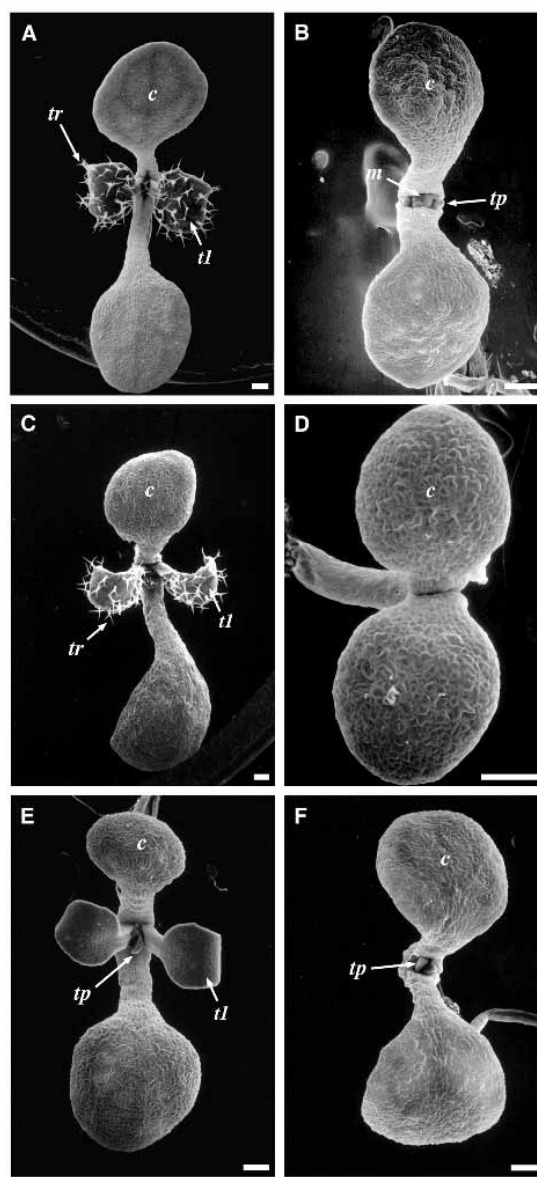


FIGURE 1.5 WILD-TYPE AND *ADD* MUTANT SEEDLING DEVELOPMENT AFTER GERMINATION AND GROWTH AT 29°C FOR 6 DAYS (PICKETT ET AL., 1996). (A) Wild-type Wassilewskija seedling with cotyledons and two true leaves, (B) the apical region of an *arrested development 1* (*add1*) seedling lacking true leaves, (C) wild-type Landsberg erecta seedling with an epicotyl displaying the first 2 true leaves. Both leaves have developed leaf blades and trichomes, (D) *arrested development 2* seedling with cotyledons, true leaf primordia are not visible, (E) wild-type Columbia glabrous1 (*gl1*) seedling with 3 true leaves. The *gl1* mutation prevents trichome development in this line; however, all true leaves have acquired dorsoventrality, (F) the apical region of an *arrested development 3* mutant. Epicotyl development is delayed in this mutant because leaf blade expansion has not occurred. However, primordial leaves are asymmetric and blade development is apparent, suggesting that these leaves have acquired dorsoventrality. Scale bar=250 mm. c, cotyledon; tp, true leaf primordium; tl, leaf from first set of true leaves; tr, trichome; m, meristematic region.

that proximity to a functional meristem is required for normal leaf morphogenesis.

The *add3* mutation does not have this effect but prevents the expansion of leaf blades at high temperature. In this mutant the temperature-dependent arrest of epicotyl development is due to a failure of normal leaf development rather than new leaf initiation.

Although all *add* mutants show a reduced rate of root growth in comparison to WT plants, temperature-dependent arrest of root growth is not observed. Given the complicated physiological interplay between shoot and root, the available data support a specific role for the *ADD* genes in regulating aspects of shoot meristem function or leaf development.

A different kind of *ts* phenotype, which involves shoot development, was observed in homozygous *topless-1* (*tpl-1*) mutants (Long et al., 2002). At the restrictive temperature the apical pole of the embryo develops as a root instead of a shoot. The result is a seedling with two roots joined at the base without hypocotyls, cotyledons or shoot apical meristem. This mutation could provide an entrance to molecular genetic study of this important, but previously experimentally intractable, decision in plant development.

FLOWER DEVELOPMENT

Ts *Arabidopsis* mutants have been useful in studying the function of floral homeotic genes, which specify the identity of floral organs. There are three classes of floral homeotic genes, designated A, B, and C (Bowman et al., 1991; Coen and Meyerowitz, 1991). Class A alone specifies sepals, A and B together cause petal formation, B and C combined specify stamens and C alone directs carpel formation. With minor variations, the ABC model has been validated in different species, including *Arabidopsis*

thaliana, *Antirrhinum majus* and *Petunia hybrida*. In *Arabidopsis*, function B is performed by a pair of homeotic genes: *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) (Goto and Meyerowitz, 1994; Hill and Lord, 1989; Jack et al., 1992). The *AP3* and *PI* proteins belong to the MADS box family of transcription factors and bind DNA *in vitro* only as *AP3/PI* heterodimers, which may explain why mutations in either class B gene cause very similar phenotypes and why the ectopic B function requires combined expression of both *AP3* and *PI* (Jack et al., 1994; Krizek and Meyerowitz, 1996).

Ts mutants have been isolated for *AP3* (*ap3-1* allele) and *DEF* (*def-101* allele) (Bowman et al., 1989). Mutant plants grown at low temperature produce nearly WT flowers, but growth at the restrictive temperature causes the development of sepals (only A function present) and the replacement of stamens by carpelloid organs (only C function present). Such conditional mutants have been useful to verify that B homeotic function is required throughout organ development and to establish the stage in organ development up to which organ identity can be redirected by activation of a different homeotic function.

Similar results were obtained using *ts* mutations in the *DEFICIENS* (*DEF*) (ortholog of *APETALA3*) and the *GLOBOSA* (*GLO*) (ortholog of *PISTILLATA*) gene of *Antirrhinum majus* (Schwarz-Sommer et al., 1992; Sommer et al., 1990; Tröbner et al., 1992; Zachgo et al., 1995).

1.2.6 AS FUNCTIONS AND INTERACTIONS OF IDENTIFIED PROTEINS COME UNDER CLOSE SCRUTINY, IT BECOMES MORE IMPORTANT TO UNDERSTAND WHICH PROPERTIES OF MUTANT PROTEINS CAUSE TEMPERATURE-SENSITIVE DEFECTS

Both heat-sensitive (not able to grow at a higher restrictive temperature) and cold-sensitive (not able to grow at a

lower restrictive temperature) mutations have been isolated in a variety of organisms. Heat-sensitive mutations, commonly referred to as ts mutations, are generally due to production of a thermolabile protein or to temperature-sensitive protein synthesis, folding or, in the case of multimeric molecules, assembly. Cold-sensitive mutations are usually found in proteins, which participate in macromolecular complexes such as ribosomes or spliceosomes.

1.2.6.1 THERE ARE TWO GENERAL CLASSES OF TEMPERATURE-SENSITIVE MUTATIONS: THOSE GENERATING THERMO-LABILE PROTEINS AND THOSE GENERATING DEFECTS IN PROTEIN SYNTHESIS, FOLDING OR ASSEMBLY

Ts mutations were initially divided into two classes by Sadler and Novick (1965).

The first class was labeled 'TL' (thermolabile) and represented mutants in which the native protein is destabilized as a consequence of the temperature-sensitive amino acid substitution (Goldenberg, 1988). Examples of this type are ts mutants of lambda repressor (Hecht et al., 1984), yeast mitochondrial hsp60 (heat shock protein 60) (Martin et al., 1992) and LAR, a human receptor-like protein tyrosine phosphatase (Tsai et al., 1991). Such mutants typically show phenotypic defects when grown first at the permissive temperature for a certain period and transferred subsequently to the restrictive temperature. The basis of 'TL' mutations lies in the heat-sensitive nature of all proteins. When heated proteins denature, their tertiary folding structure is transformed into a variable, random-coil form. Protein folding is a process, which is accompanied by an unfavourable decrease in entropy and thus, driven by enthalpy. However, elevated temperatures increase the entropy term in the free energy equation (ΔG (free energy difference) = ΔH (enthalpy difference) – T

(temperature) . ΔS (entropy difference)), and hence favour the random-coil form. Each protein has a specific temperature at which it unfolds, called the transition temperature. So, 'TL' mutations are missense mutations that result in a destabilized protein with a lower transition temperature. If the transition temperature falls within the normal growth range of the organism, then it is possible to grow the mutant organism at a low temperature, where the protein is folded and active, and at a high temperature, where the protein is unfolded and inactive.

The second class of ts mutations was labeled 'TSS' (temperature-sensitive synthesis). These mutations, like the 'TL' class, are defective when mutants are grown at the restrictive temperature. However, if mutants are grown at the permissive temperature, and then shifted to the restrictive temperature, they do not display phenotypic perturbances. The deficiency only occurs if the mutant protein was expressed and permitted to assemble at the restrictive temperature. The defects in gene product formation caused by 'TSS' mutations are varied. Most of the ts mutants of this type are defective in protein synthesis, folding or assembly. However, Singer and Gold (1976) described a 'TSS' mutation in the rII gene of phage T4 that probably acted at the level of messenger RNA structure or function and ts mutants of tRNAs exist (Smith et al., 1970). Mutants of this class include ts mutants of more than 30 genes of bacteriophage T4 (Edgar and Lielausis, 1964), ts mutants of T4 DNA polymerase (Dewaard et al., 1965) and ts alleles at more than 30 sites in the P22 tailspike endorhamnosidase (Smith and King, 1981).

A systematic examination of 'TSS' mutations in the tailspike protein of phage P22 revealed that the primary defect was in the folding of the polypeptide chain (Goldenberg and King, 1981; Goldenberg et al., 1983; Sturtevant et al., 1989; Mitraki et al., 1991). The temperature-sensitive

amino acid substitutions had no effect on the stability or activity of the thermostable tailspike protein, once matured at the permissive temperature. However, if expressed at the restrictive temperature, a folding intermediate was unable to continue productive folding. This subclass of 'TSS' mutations was subsequently called 'TSF' (temperature-sensitive folding) (Yu and King, 1984; King et al., 1989). Similar 'TSF' phenotypes have also been reported for mutants of D-lactate dehydrogenase (Truong et al., 1991) and bacterial luciferase (Sugihara and Baldwin, 1988).

THE ABSENCE OF INTRAGENIC COMPLEMENTATION IS A GENERAL PROPERTY OF 'TSF' (TEMPERATURE-SENSITIVE FOLDING) MUTATIONS IN GENES ENCODING FOR MULTIMERIC PROTEINS

Many sets of missense mutants, including *ts* mutants, exhibit intragenic complementation. This kind of complementation generally occurs in genes encoding different subunits of multimeric proteins and therefore provides a window into polypeptide chain domains and subunit interactions (Fincham, 1966).

Two general models have been proposed to explain intragenic complementation. First, if each allele has knocked out a different enzymatic function of the protein, the presence of both alleles may restore full protein activity (McGavin, 1968). Second, if the protein is oligomeric, the two alleles may help each other to correctly fold and assemble (Crick and Orgel, 1964).

Interaction of missense polypeptide chains, following this second model, has been demonstrated for many different oligomeric proteins (Schlesinger and Levinthal, 1963; Zabin and Villarejo, 1975). Significant intragenic complementation has been reported for *ts* mutations within the genes for a number of phage T4 structural proteins (Bernstein et al., 1965) and *ts* mutations in the SV40 major coat protein (Behm et al., 1988). Therefore, it is assumed that direct

(subunit) interaction corrects the mutant defects and hence provides the mechanism for intragenic complementation.

However, Gordon and King (1994) found that, despite the intimate interactions between coat protein molecules of phage P22 during the subunit assembly process, 'TSF' coat protein mutants did not show any intragenic complementation. They were also recessive. These genetic properties reflect the nature of the intracellular defect: defective folding of the polypeptide chain. The same recessive character and lack of intragenic complementation has been found with 'TSF' mutants of the trimeric P22 tailspike protein (Smith et al., 1980; Goldenberg et al., 1983). These properties are consistent with the idea that the mutant coat protein molecules are not defective in the stability (or function) of the folded molecule, but of an assembly-incompetent precursor. Subunit/subunit interaction between different alleles is expected to offer the possibility of improving the stability or function of folded assembly-competent molecules, but not of destabilized 'TSF' folding intermediates.

Therefore, a 'TSF' defect can be defined as the destabilization of a folding intermediate at the restrictive temperature, which prevents the mutant chain from reaching the conformation required for subunit/subunit recognition.

SELECTION OF SECOND-SITE SUPPRESSORS OF TEMPERATURE-SENSITIVE MUTANTS HAS PROVIDED A GENERAL APPROACH FOR IDENTIFYING GENES OF WHICH PRODUCTS INTERACT WITH AN INITIAL PROTEIN OF INTEREST

For conditional-lethal mutants, restrictive conditions provide an absolute selection for further mutation, since only revertant individuals can grow. Such revertants need not be true WT, rather, they may have acquired suppressors, new mutations that act in a way that corrects, replaces, or bypasses the original defect.

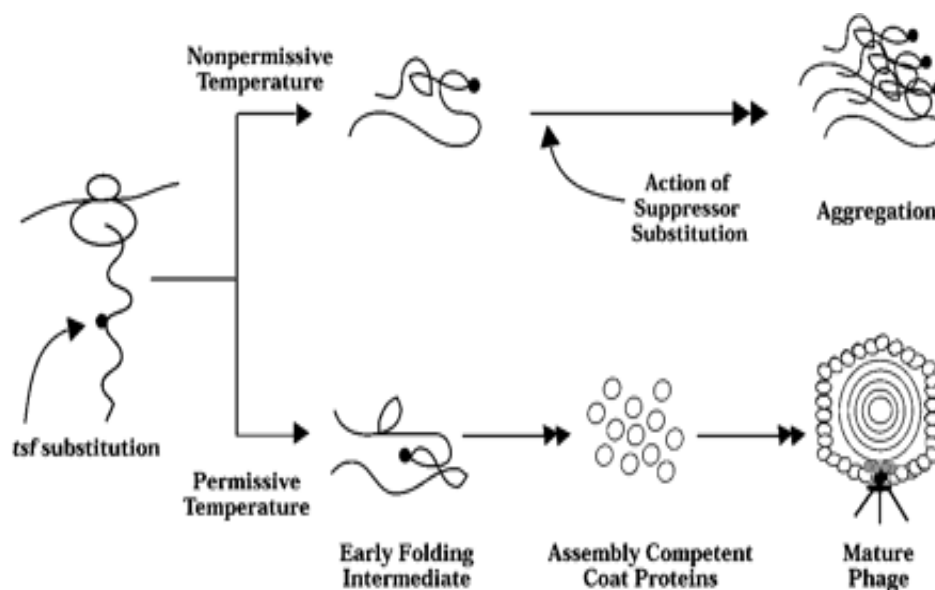


FIGURE 1.6 MODEL OF THE FOLDING PATHWAY OF P22 COAT PROTEIN CARRYING 'TSF' AMINO ACID SUBSTITUTIONS (ARAMLI AND TESCHKE, 1999). A newly synthesized polypeptide forms an intermediate that at permissive temperature proceeds to an assembly-competent coat monomer that is necessary to form a capsid. At a nonpermissive temperature, the tsf amino acid substitution results in an intermediate that is aggregation-prone and forms inclusion bodies. A suppressor amino acid substitution could prevent 'TSF' action.

The distinction between 'TSF' and ts mutations that perturb the function of the already folded protein ('TL' (thermolabile) or assembly-defective 'TSS' (temperature-sensitive synthesis)) is likely to be particularly important in the isolation of second-site suppressor mutations (figure 1.6). If the initial mutation impairs the function of the active intracellular conformation ('TL' or assembly-defective 'TSS'), gene products interacting with the mature form of the protein can be isolated. However, if the initial mutation destabilizes a folding intermediate such selection may yield sites influencing folding and maturation pathways (Fane and King, 1991).

Global suppressors are a type of second-site suppressors that counter the deleterious effects of a wide range of protein folding and stability mutations. So, the characterization of mutants that alter folding and stability of proteins may help identify residues that specify critical steps in the folding pathway or ones that highlight key interactions in the native

structure. In a selection for second-site suppressors of tailspike 'TSF' mutations global intragenic suppressors were obtained (Fane et al., 1991; Mitraki et al., 1991). By suppressing the pathway to inclusion body formation, these substitutions improved the folding efficiency of different 'TSF' mutations, as well as of the WT tailspike protein.

1.2.6.2 COLD-SENSITIVE MUTATIONS INVOLVE AN ALTERED PROTEIN CONFORMATION THAT IS STILL ABLE TO INTERACT AT HIGHER TEMPERATURES BUT FAILS TO DO SO AT LOWER TEMPERATURES

Cold-sensitive mutations affect the ability of proteins to form multimeric or macromolecular complexes. Despite their altered conformation as a consequence of an amino acid change, these proteins are able to function and interact at higher temperature. In fact, due to the increased kinetic activity, the complex can adjust to accommodate the mutant protein. However, at lower temperatures the

altered shape and slower movement of the mutant protein prevents it from interacting efficiently (O'Donovan and Ingraham, 1965).

The remarkable property of self-assembly of ribosomal proteins is highly dependent upon temperature, an observation that led Nomura to the brilliant decision to screen cold-sensitive lethals of *Escherichia coli* for defects in ribosomal proteins (Guthrie et al., 1969). It was found that a significant proportion of cold-sensitive lethal mutations in this organism do indeed encode defective ribosomal proteins.

Although cold-sensitive mutations are isolated to a lesser extent, they are included in several other specific studies such as sporulation (Esposito and Esposito, 1968) and mitochondrial petite (Weisblum and Davies, 1968) in yeast and developmental mutants in *Drosophila* (Foster and Suzuki, 1970).

1.2.6.3 TEMPERATURE- AND COLD-SENSITIVE MUTATIONS DIFFER STRONGLY IN GENETIC DISTRIBUTION AND TARGET GENES

Edgar and Lielausis (1964) found that ts mutations were widely distributed over the genome of bacteriophage T4, although they were not distributed completely randomly. They observed long segments in which they didn't find any ts mutations, while other regions were densely populated with such mutations. However, they showed that modifications of the isolation procedures did yield ts mutations located in otherwise sparsely populated segments of the genome. Thus, it would appear probable that many of the blank segments were regions of low detectable mutability rather than regions in which ts mutations cannot occur. Also in *Drosophila* it was observed that, given a proper screening protocol, temperature-sensitivity is a property that can be readily detected for a wide spectrum of loci throughout the genome (Suzuki, 1970). As

a consequence, ts lethals are defective with respect to a wide range of functions.

In contrast, cold-sensitive mutants are remarkably clustered in a small number of cistrons (Scotti, 1968; Cox and Strack, 1970) and involve changes primarily in regions of proteins concerned in regulatory functions. Striking is the case of ribosome assembly in *Escherichia coli*, where extensive searches for ts mutants were fruitless while cold-sensitive mutants were readily obtained (Tai et al., 1969).

1.2.7 DIFFERENT STRATEGIES CAN BE FOLLOWED TO GENERATE TEMPERATURE-SENSITIVE MUTANTS

1.2.7.1 TEMPERATURE-SENSITIVE MUTANTS ARE TYPICALLY GENERATED BY RANDOM MUTAGENESIS FOLLOWED BY LABORIOUS SCREENING PROCEDURES

Ts mutants are usually point mutations, which are often generated by random mutagenesis followed by screening of large numbers of progeny. In most of the organisms studied, the highest frequency (25-50%) is observed when chemical mutagenesis is applied using the alkylating agent EMS (ethyl methane sulfonate). γ -ray-induced ts mutations show a frequency of 3%-3.5%. In the case of prokaryotes and yeast, such procedures work well because simple screens or selection methods exist and a large number of progeny can be simultaneously screened using simple plate assays. In more complicated organisms, however, such an approach suffers from several drawbacks. Since mutations will be generated throughout the genome, a large number of progeny need to be screened before a ts mutant is obtained. In the case of *Drosophila melanogaster* and *Arabidopsis thaliana*, this number is typically of the order of several hundred thousand. Screening such a large number of progeny can be extremely laborious in situations where a simple screen does not

exist. Furthermore, in organisms with long generation times or in cases where it is not possible to obtain large numbers of progeny, random mutagenesis of the entire genome cannot be used to isolate ts mutants.

Identifying and recovering ts mutations is also hindered by their rarity, a particular problem in higher organisms. It is also clear that ts alleles cannot be generated for all proteins and some remain refractory even following extensive mutagenesis (Harris et al., 1992).

1.2.7.2 HEAT-INDUCIBLE DEGRON: AN ALTERNATIVE METHOD FOR CONSTRUCTING TEMPERATURE-SENSITIVE MUTANTS

Features of proteins that confer metabolic instability are called degradation signals, or degrons (Varshavsky, 1991). The essential component of one degradation signal, the first to be discovered, is a destabilizing N-terminal residue of a protein (Bachmair et al., 1986; Varshavsky et al., 1992, 1996a). This signal is called the N-degron. The collection of N-degrons containing different destabilizing residues yields a rule, termed the N-end rule of protein degradation, which relates the *in vivo* half-life of a protein to the identity of its N-terminal residue. The N-end rule pathway is present in all organisms examined, including the bacterium *Escherichia coli* (Tobias et al., 1991; Shrader et al., 1993), the yeast *Saccharomyces cerevisiae* (figure 1.7; Bachmair and Varshavsky, 1989) and mammalian cells (Gonda et al., 1989; Lévy et al., 1996). In eukaryotes, the N-degron consists of at least two determinants: a destabilizing N-terminal residue and a specific internal lysine residue (or residues) of a substrate. The lysine residue is the site of attachment of a multiubiquitin chain. Ubiquitin is a 76-residue protein whose covalent conjugation to other proteins plays a role in a number of cellular processes, primarily

through routes that involve protein degradation.

Dohmen et al. (1994) devised a method for constructing ts mutants, which involves making fusions of ubiquitin-ts-arginin dihydrofolate reductase to the protein of interest (figure 1.8). Arginin-dihydrofolate reductase (Arg-DHFR) bearing an amino-terminal arginin is long-lived in the yeast *Saccharomyces cerevisiae*, even though arginin is a destabilizing residue in the N-end rule of protein degradation. A temperature-sensitive derivative of Arg-DHFR (Arg-DHFR-ts) was identified that is long-lived at 23 °C but rapidly degraded by the amino-end rule pathway at 37 °C. Fusions of ts Arg-DHFR to either Ura3 (orotidine 5'-phosphate decarboxylase) or Cdc28 (the main cyclin-dependent kinase involved in cell cycle control in budding yeast) of *Saccharomyces cerevisiae* confer ts phenotypes specific for these gene products. Thus, Arg-DHFR-ts is a heat-inducible degradation signal that can be used to produce ts mutants without a mutational screening.

Although this is a powerful and elegant approach, it remains to be seen how well it works in other organisms. In some cases, it is possible that the fusion partners may also affect the normal functioning of the protein at temperatures below the restrictive temperature.

1.2.7.3 FOR GLOBULAR PROTEINS TEMPERATURE-SENSITIVE MUTANTS CAN BE PREDICTED AND DESIGNED SOLELY BASED ON THE AMINO ACID SEQUENCE

A method to predict those residues in a protein sequence that, when appropriately mutated, are most likely to give rise to a ts phenotype was described by Varadarajan et al. (1996).

Their approach was based on the fact that mutations at buried residue positions cause much larger changes in protein stability than mutations at surface positions (Pakula et al., 1986; Kellis et al., 1989; Shortle et al., 1990). Decreased

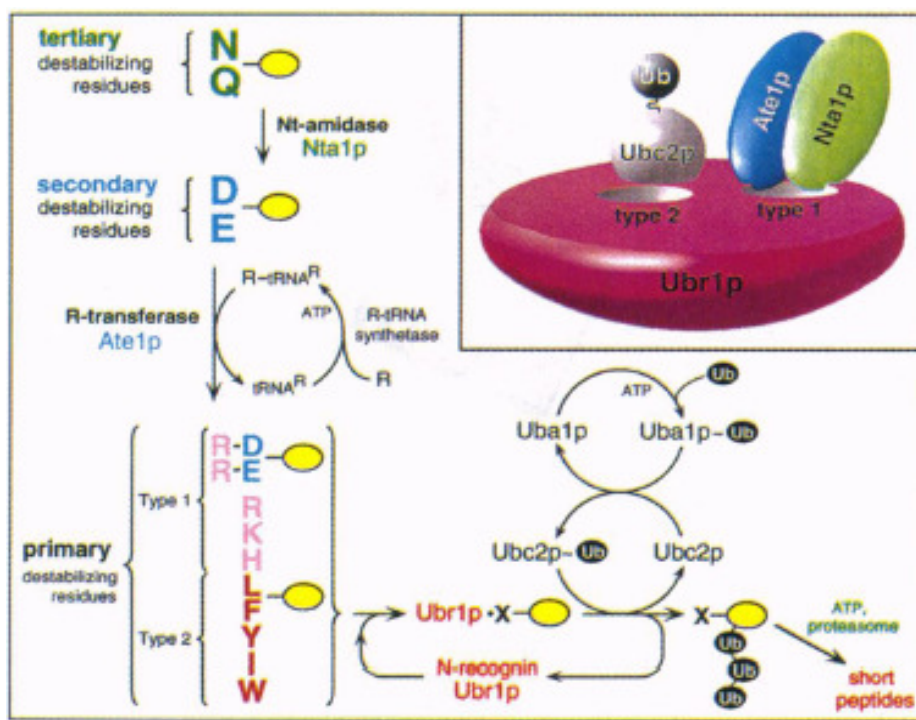


FIGURE 1.7 THE *SACCHAROMYCES CEREVISIAE* N-END RULE PATHWAY (VARSHAVSKY, 1996A). Type 1 (N-d^{P1}) and type 2 (N-d^{P2}) primary destabilizing N-terminal residues are in purple and red, respectively. Secondary (N-d^S) and tertiary (N-d^T) destabilizing N-terminal residues are in blue and green, respectively. The yellow ovals denote the rest of a protein substrate. The conversion of N-d^T residues N and Q into N-d^S residues D and E is mediated by N-terminal amidohydrolase (Nt-amidase), encoded by *NTA1*. The conjugation of the N-d^{P1} residue R to N-d^S residues D and E is mediated by Arg-tRNA-protein transferase (R-transferase), encoded by *ATE1*. A complex of N-recogin and the ubiquitin (Ub)-conjugating (E2) enzyme Ubc2p catalyzes the conjugation of activated Ub, produced by the Ub-activating (E1) enzyme Uba1p, to a lysin residue of the substrate, yielding a substrate-linked multi-Ub chain. Uba1p-Ub and Ubc2p-Ub denote covalent (thioester-mediated) complexes of these enzymes with Ub. A multiubiquitylated substrate is degraded by the 26S proteasome. Inset: A model of the targeting complex. The 20-kDa Ubc2p E2 enzyme is depicted carrying activated Ub linked to cystein-88 of Ubc2p through a thioester bond. Both the 52-kDa Nta1p (Nt-amidase) and the 58-kDa Ate1p (R-transferase) bind to the 225-kDa Ubr1p (N-recogin) in proximity to the type 1 substrate-binding site of Ubr1p. In addition, Nta1p directly interacts with Ate1p.

protein stability is correlated with reduction of the protein levels *in vivo* and with the generation of a ts phenotype (Parsell et al., 1989; Pakula et al., 1989). The strategy has therefore been to predict those residues in the sequences that are likely to be buried in the protein structure. Values of two sequence-based parameters, the average hydrophobicity (Rose and Roy, 1980; Hopp and Woods, 1981) and the hydrophobic moment (Eisenberg, 1984), were used for the prediction. The average hydrophobicity is calculated over a seven-residue window centered around the residue of interest. The hydrophobic

moment is a vectorial sum of hydrophobicities calculated over a nine-residue window, with a phase angle optimized for detection of amphiphilic helical structures, which typically would not have a high hydrophobicity. Stringent cut-offs for these two parameters ensured that it was possible to predict buried hydrophobic residues with an accuracy >80%. The method successfully predicted several known ts and partially active mutants of T4 lysozyme, λ repressor, gene V protein and staphylococcal nuclease. Furthermore, it correctly predicts residues that form part of the hydrophobic cores of

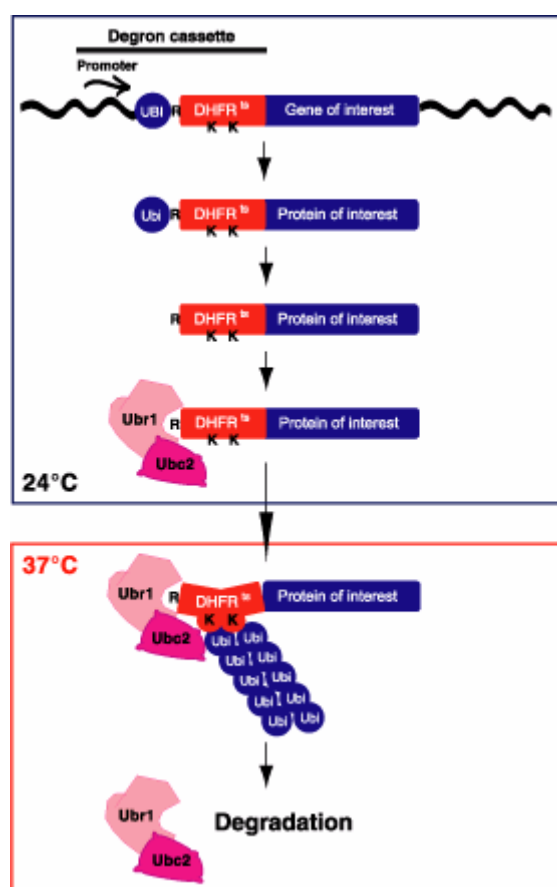


FIGURE 1.8 THE HEAT-INDUCIBLE DEGRON CAUSES A PROTEIN TO BE DEGRADED AT 37°C (SANCHEZ-DIAZ ET AL., 2004).

λ repressor, myoglobin and cytochrome b562.

It should be noted that the authors did not attempt to predict all possible ts mutants. Ts mutations can result from substitutions at non-core residues and at surface positions that may be important for folding, proteolytic sensitivity or interaction with other proteins. Rather, the method is designed to yield ts mutants at sufficiently high frequencies to be experimentally useful in systems in which it is not feasible to screen large numbers of progeny generated by random mutagenesis.

In a more recent report Chakshumathi et al. (2004) tested this

algorithm to further refine and improve the method. Initially, they used *Escherichia coli* cytotoxin CcdB as a model system and developed simple procedures for generating ts mutants at high frequency through site-directed mutagenesis. Putative buried, hydrophobic residues are selected through analysis of the protein sequence, as described above. Residue burial is then confirmed by ensuring that substitution of the residue by asparagin leads to protein inactivation. At such sites, a ts phenotype can typically be generated either by substitution of two predicted, buried residues with the remaining 18 amino acids or by introduction of lysin, serin, alanin and tryptophan at three to four predicted buried sites. By using these design strategies, 17 tight ts mutants of CcdB were isolated at four predicted buried sites. The rules were further validated by making several ts mutants of the yeast transcriptional activator Gal4 at residues 68, 69 and 70. Gal4 binds to a sequence called UAS (upstream activating sequence consisting of tandem 17 base pairs, imperfect repeats) and activates transcription of genes downstream of this sequence (Bram and Kornberg, 1985; Fedor et al., 1988; Giniger et al., 1985). Gal4 is absent in higher eukaryotes. However, if expressed under control of a tissue- or developmental stage-specific promoter, the gene placed downstream of UAS will be expressed in a Gal4-dependent pattern. The UAS-Gal4 system is widely used in cell culture and in a variety of organisms, including yeast, fruitflies, zebrafish, mice and frogs (Andrulis et al., 1998; Bernard et al., 1993; Brand and Perrimon, 1993; Hartley et al., 2002; Ornitz et al., 1991; Rahner et al., 1996; Scheer and Camnos-Orthega; 1999). No ts mutants of Gal4 have currently been reported. Such mutants would be very useful, because they would permit reversible, conditional expression of the very large number of UAS constructs that have already been made.

1.2.7.4 TEMPERATURE-SENSITIVE MUTANT INTEINS INSERTED WITHIN TWO TRANSCRIPTION FACTORS PROVIDE A NEW TOOL FOR CONTROLLING GENE EXPRESSION IN YEAST AND FLIES

Zeidler et al. (2004) combined ts inteins - protein elements capable of self-excision from a precursor protein and ligation of the flanking protein regions - with the yeast transcription factors Gal4 and Gal80 to create a system that overcomes the problems associated with the isolation of ts mutants using random mutagenesis (figure 1.9).

Intein-mediated protein splicing is a post-translational, autocatalytic reaction requiring no known host cell cofactors, making inteins very portable. However, each intein is more or less sensitive to proximal host protein sequences, which are the substrate of the intein enzymatic activity (Chong et al., 1998; Noren et al., 2000; Southworth et al., 1999). Over 170 inteins have been catalogued in InBase, the intein database (Perler, 2002). In naturally occurring precursors, protein splicing is usually required for host protein activity. However, when inserting inteins into new sites, host protein activity of the precursor and the spliced product should be tested using inactive and active inteins, respectively. The positional effects of protein splicing in foreign contexts can often be overcome by trying several intein insertion sites. Zeidler et al. have eliminated this problem by inserting temperature-dependent splicing mutants of the *Saccharomyces cerevisiae* vacuolar ATPase subunit (VMA) intein into transcriptional regulators that control expression of essentially any target protein.

The transcriptional regulators they chose to modify were Gal4 and Gal80. Gal80 functions by binding to Gal4 and blocking the latter's ability to activate transcription. Zeidler et al. demonstrated intein-mediated conditional target gene expression in yeast, *Drosophila melanogaster* tissue culture cells and in transgenic *Drosophila melanogaster* imaginal discs (cell cluster progenitors of adult organs). This method is readily transferable to the

multitude of genes currently regulated by Gal4.

They initially genetically selected temperature-dependent splicing mutants of the *S. cerevisiae* VMA intein in Gal4 (Gal4INT^{ts}) in yeast. These mutant inteins were then transferred to two sites in Gal80 (Gal80INT^{ts}). Efficient, temperature-dependent splicing of both Gal80INT^{ts} constructs was observed in yeast, but only one Gal80INT^{ts} fusion spliced efficiently in *D. melanogaster* cells, reflecting the complexity of generating a universal ts intein. This is possibly the first example of transfer of ts splicing to a second host protein, indicating that the ts phenotype is intrinsic to the intein.

Transcription of a target gene is stimulated when Gal4 binds to an upstream activator sequence (UAS) inserted 5' to the gene (see also §1.2.7.3). A second layer of control is added by Gal80, which represses transcription of the target gene by binding to Gal4. When the ts intein is present in Gal4, target gene expression is activated only at the permissive-splicing temperature. When the ts intein is present in Gal80, Gal4 activity is only repressed at the permissive-splicing temperature.

Using this system, any Gal4 line can be directly converted into a conditional expression system by merely crossing with a Gal80INT^{ts} or Gal4INT^{ts} line. The combination of available temporal-, tissue- and stage-specific expression pattern drivers provide a multitude of combinations of Gal4 and Gal80INT^{ts} interactions, resulting in an enormous diversity of potential target protein expression patterns. The ability to activate Gal80INT^{ts} by shifting to the permissive-splicing temperature provides another layer of control.

The elegance of this system was demonstrated in the transgenic *Drosophila melanogaster* experiments. This is the first example of intein-mediated protein splicing in a multicellular animal (as opposed to splicing in plants or pathogenic bacteria, viruses and yeast that invade complex organisms).

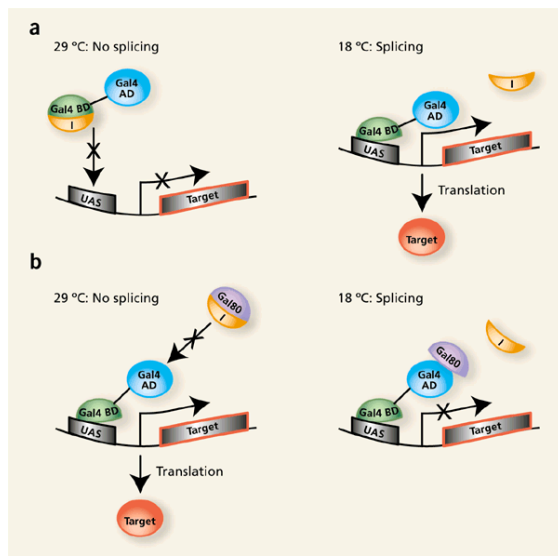


FIGURE 1.9 **CONDITIONAL TARGET GENE EXPRESSION CONTROLLED BY TEMPERATURE-SENSITIVE SPLICING OF INTEIN FUSIONS** (ZEIDLER ET AL., 2004). Gal4 is a two-domain protein that activates transcription of genes preceded by an upstream activator sequence (UAS). The Gal4 N-terminal binding domain (BD) binds to UAS, whereas the carboxy-terminal activation domain (AD) stimulates transcription of the target gene. Gal80 represses Gal4 by binding to the Gal4 activation domain. **(a)** A temperature-sensitive splicing mutant of the *S. cerevisiae* VMA intein (I) was cloned into Gal4. Spliced Gal4 is fully functional, whereas the Gal4INT^{ts} precursor fails to activate transcription of the target gene. **(b)** A temperature-sensitive splicing mutant of the *S. cerevisiae* VMA intein (I) was cloned into Gal80. The Gal80INT^{ts} precursor has no effect on Gal4, whereas spliced Gal80 binds to the transcription activation domain (AD) of Gal4 repressing transcription of the target gene.

Expression of UAS-green fluorescent protein (GFP) reflected the spatial location and temperature-dependent splicing of *apterous* enhancer driven Gal4INT^{ts}. In a second experiment, Gal80INT^{ts} was under the control of the *decapentaplegic* enhancer and Gal4 was under control of the *patched* enhancer, yielding different spatial patterns of expression of the repressor and activator in imaginal discs. Gal4 activation of UAS-GFP and UAS-Notch was dependent on the spatial location of both Gal4 and Gal80, and on temperature-dependent splicing of Gal80INT^{ts}.

The ts intein system is directly applicable to any organism that uses Gal4 activation. It is limited to cells or organisms that can undergo the requisite temperature shift, making it impractical in

many warm-blooded animals. However, other ways of controlling protein splicing exist or are in development. These include splitting the precursor and splicing *in trans* (Noren et al., 2000) or using small molecules to activate or inhibit splicing (Mootz and Muir, 2002). Several genetic selection systems are now available to screen for ts inteins and protein splicing inhibitors and activators (Paulus, 2003).

One of the most significant contributions of the authors is the concept of controlling a universal regulator with ts inteins. The plethora of Gal4 and UAS lines currently available and easily adapted to this system makes the conditional splicing of Gal4INT^{ts} and Gal80INT^{ts} extremely important new genetic tools. They have demonstrated conditional protein splicing in a complex eukaryote and have identified ts mutants that could possibly be transferred to other host proteins, providing a framework for making other ts regulator proteins.

Although the above-described approaches offer elegant alternatives to overcome the problems associated with random mutagenesis, the latter is still the most widely used method to induce (possibly ts) point mutations. Despite the mentioned drawbacks, it is a technically and experimentally simple and effective procedure that has been used for more than 60 years now in all kinds of organisms.

1.3 ESTABLISHING GENE FUNCTION THROUGH MUTAGENESIS

Gene function can be revealed through analysis of mutants by either forward or reverse genetics. The methodologies that are currently available to obtain mutants with desired phenotypes or mutations in a gene of interest and to clone a mutant gene will be discussed in detail for the model plant *Arabidopsis thaliana*.

1.3.1 *ARABIDOPSIS THALIANA* AS AN EXPERIMENTAL SYSTEM FOR MOLECULAR GENETIC STUDIES

The small mustard weed *Arabidopsis thaliana* (family Brassicaceae) was known to botanists for at least four centuries and had been used in experimental research for about half a century, before Friedrich Laibach outlined the advantages of using it in genetic experiments (figure 1.10; Laibach, 1943; Rédei, 1992). In 1943 he also suggested that it could be used as a plant model system. Several features make *Arabidopsis* amenable to classical experimental genetics: a small size, a rapid generation time (5-6 weeks under optimum conditions), the ability to grow well under controlled conditions (either in soil or in defined media), high fecundity (up to 10,000 seeds per plant) and the ease with which a mutant line can be maintained (by self-fertilization) and out-crossed.

As well as having all the key features of a model system for classical genetics, *Arabidopsis* has the smallest known genome among flowering plants (125Mb), with fewer repetitive sequences than any known higher plant, greatly facilitating molecular studies and map-

based cloning (Leutwiler et al., 1984; Pruitt and Meyerowitz, 1986; The *Arabidopsis* Genome Initiative, 2000). In addition, *Arabidopsis* can easily be transformed by *Agrobacterium*-mediated gene transfer, which is a prerequisite for many molecular genetic experiments. The advantages of using *Arabidopsis* in such experiments have been reviewed extensively (Meinke et al., 1998; Meyerowitz, 1989; Meyerowitz and Pruitt, 1985). Although *Arabidopsis* had been used for classical genetic studies for a long time, its role was not comparable with that of *Drosophila* in the systematic analysis of developmental and metabolic processes. Only once it was recognized that *Arabidopsis* had a small genome and, therefore, the potential to combine classical genetics with molecular studies, did it advance to being the most studied higher plant. The success of *Arabidopsis* as a model organism is mainly due to its amenability to forward genetic screens, by which genetic variation is artificially induced and mutagenized plants are screened for phenotypes of interest. Mutagenesis in *Arabidopsis* mainly relies on ethylmethane sulphonate (EMS) and, to a much lesser extent, fast neutron radiation (FNR). The study culminates in the molecular isolation of the mutated gene, which causes the observed phenotype. So, a function can be assigned to a molecular change in the inherited material of an individual. As more scientists joined the *Arabidopsis* community, experience was gained and tools were developed to carry out these screens with the aim of isolating genes that are involved in all aspects of plant life.

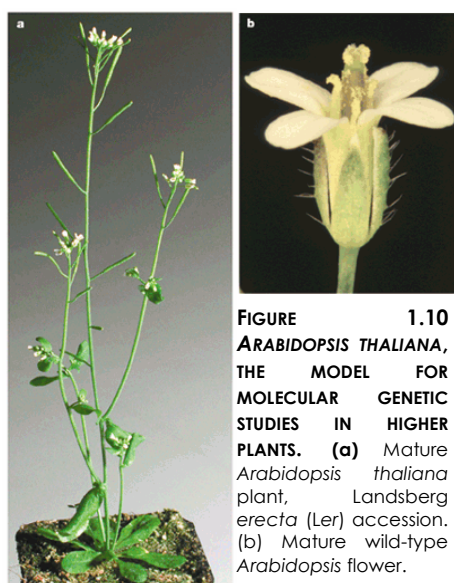


FIGURE 1.10
***ARABIDOPSIS THALIANA*,
THE MODEL FOR
MOLECULAR GENETIC
STUDIES IN HIGHER
PLANTS. (a)** Mature
Arabidopsis thaliana
plant, Landsberg
erecta (Ler) accession.
(b) Mature wild-type
Arabidopsis flower.

1.3.2 FORWARD GENETICS: FROM PHENOTYPE TO GENOTYPE

Forward genetics aims to identify the sequence change that underlies a specific mutant phenotype. The mutant phenotype can be the result of deliberate mutagenesis or be defined on the basis of existing variation. When the mutation

causing the phenotype is the result of a T-DNA or transposon insertion, rapid identification of the gene of interest is at least theoretically possible by locating the sequence tag and analyzing its neighboring sequences. However, the identification of a gene affected by a chemically or radiation-induced mutation requires a more laborious map-based cloning approach in which markers linked to the mutated gene are used to delimit the region containing the gene of interest (Østergaard and Yanofsky, 2004).

1.3.2.1 MUTAGENESIS IN *ARABIDOPSIS*

In order to conduct genetic analysis, at least two alleles for a given locus must exist and the facile production of these alleles through mutagenesis has been the focus of considerable research since the early part of the 20th century (Koorneef, 2002).

CLASSICAL MUTAGENESIS

Most of the mutations available in *Arabidopsis thaliana* have been induced by some form of ionizing radiation or by chemical agents. Since the frequency of insertional mutagenesis is generally lower, it may be advantageous to first define a locus by these classical methods. Once one knows that a particular kind of mutation is possible, it is then practical to plan a strategy for the isolation of the corresponding gene.

The generation of plants, grown from the mutagen-treated parental genotype, is called the M₁ generation and is heterozygous for mutations. The embryos within the seeds formed on such plants and the plants that grow out of these seeds represent the M₂ generation. The M₂ is the first generation following mutagenesis in which homozygous recessive mutations can be detected. For this reason, it is the generation most frequently used in mutant screens (Feldmann et al., 1994).

Plants from Treated Seeds Are Chimeras

Seeds are the most commonly mutagenized *Arabidopsis* tissue, because of their ready availability, robustness in mutagenesis and ease of propagation (figure 1.11). Each cell within the original treated seed is mutagenized in its own unique way. Thus, different cells of the same seed will contain different mutations. The M₁ plant that sprouts from the seed is therefore a chimera of genetically different cell lineages, each of which we would expect to be heterozygous for whatever new mutations have been induced. The population of M₂ seeds from a given M₁ plant contains both heterozygotes and homozygotes for any mutation that was induced in a cell whose descendants gave rise to gametes.

Fortunately, a small number of cells in the mature seed give rise to the body of the plant and subsequently contribute to the next generation. So, only those mutations that occurred in cells that form the germ line (precursor cells of the gametes) can be detected in the M₂ generation. The number of cells that contribute to the germ line was called the genetically effective cell number (GECN) by Li and Redei (1969) and was estimated to be, on average, two in *Arabidopsis*. The consequence is that half of the progeny of an individual M₁ plant derives from a sector in which a specific gene (A) is not mutated (AA) and the other half from the sector that was heterozygous for this gene (Aa). The segregation ratio observed for aa in the progeny of this chimeric plant is then 7A (AA + Aa) : 1aa. This 7 : 1 ratio is obtained from a 4 : 0 ratio from the non-mutated sector and a 3 : 1 ratio from the sector carrying the mutated allele a. The sectors that originate from the germ line cells present in seeds are not identical in size and their contribution to the main inflorescence on which seeds are harvested, changes during development. Parts of the plant that are formed later (the top) are mostly composed of only one sector from only one germ line cell.

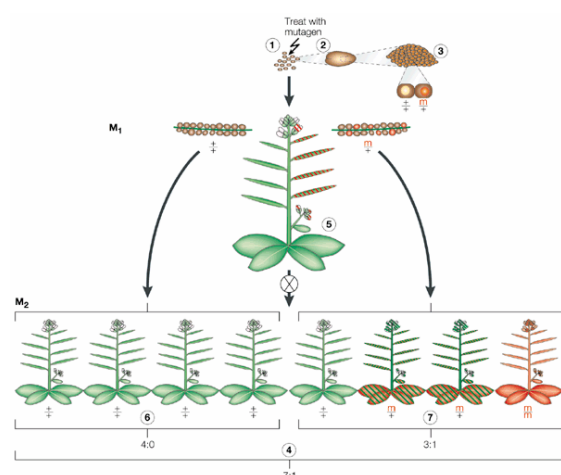


FIGURE 1.11 THE BASIS OF SEED MUTAGENESIS IN *ARABIDOPSIS* (ALBERTS ET AL., 1998). The most practical way to induce heritable mutations in *Arabidopsis* is by seed mutagenesis (1). The targets of seed mutagenesis are the diploid cells of the fully developed embryo (2) covered by the seed coat. To assess the effectiveness of mutagenesis, it is crucial to know how many of the targeted cells will eventually contribute to the next generation. Although plants do not have a determined germ line, in practice, the functional germ line can be defined as the number of cells in the shoot meristem of the embryo (3) that will contribute to the seed output (the genetically effective cell number (GECN)). The GECN is determined genetically and might vary between different species and for different developmental stages. If the segregation in the M_2 , the generation derived from mutagenized M_1 plants, is 3:1, the functional germ line consists of a single cell at the time of mutagenesis (GECN=1). In *Arabidopsis*, recessive mutants segregate in a ratio of 7:1 in an M_2 population (4), therefore, GECN = 2. Because the functional germ line consists of two cells at the time of mutagenesis, the developing M_1 (which develops from the mutagenized seed) is chimeric and consists of two sectors, which might vary in size (5). Mutations segregate 4:0 (6) or 3:1 (7), depending on the sector from which the M_2 seeds derive. To recover a potential mutation, it is therefore necessary to analyse M_2 seeds from different branches of the M_1 plant.

The consequence is that variation occurs between M_1 plants and that the harvesting policy also determines how many sectors contribute to the progeny. Chimerism disappears when the mutation is transmitted through the progeny because every individual of this progeny is derived from a single zygote cell.

To Pool or Not to Pool?

To reduce labor, M_2 seed is often collected from pools of M_1 plants, which is a good strategy for mutations that do not

cause death or sterility when homozygous. If the mutations of interest are expected to cause lethal or sterile phenotypes, then M_2 seeds should be collected from individual M_1 plants. This way, when an interesting mutant is found among the progeny of a particular M_1 , plants heterozygous for the mutation can be recovered from the siblings of the dead or sterile plant. Another advantage of collecting seeds from individual plants is that it prevents the same allele from being recovered more than once. If two plants with the same phenotype are found in the same pool, there is a good chance that they will carry the same mutation. Plants from different pools will carry independent mutations. So, the higher number of M_1 pools, the more independent alleles will be recovered.

Physical Versus Chemical Mutagens

The most commonly used physical mutagens are ionizing radiation (γ - and X-rays) and fast neutron bombardment. Each type of radiation produces deletions at a high frequency. The deletions are supposed to be larger after fast neutron radiation. The latter type of radiation yields densely ionized tracks compared to the mainly sparse ionizations caused by γ - and X-rays. For this reason fast neutrons are sometimes preferred. However, facilities where one can apply this type of radiation are rare in comparison to sources of γ - and X-rays. Deletions are attractive for many experiments because the mutants are often true nulls. In addition, a deletion is often detectable with Southern blot and PCR analysis, which facilitates the recognition of the gene to be cloned using map-based cloning. Although large deletions (parts of chromosome arms) can occur, they are rarely observed in M_2 populations because a large deletion in most cases affects the viability of germinating pollen grains that then will not contribute to fertilization. This effect of reduced transmission of mutated pollen is called certation.

Chemical mutagenesis was developed by Charlotte Auerbach and her collaborators in England prior to and during the Second World War as part of a chemical warfare program. Chemical mutagens can be divided into two classes: the alkylating agents and the cross-linking agents.

The alkylating mutagens cause high point mutational densities with only low levels of chromosome breaks that would cause aneuploidy, reduced fertility and dominant lethality. Ethyl methane sulphonate (EMS) has been the most commonly used mutagen in plants and induces large numbers of recessive mutations per genome. Other alkylating agents such as ethylnitrosourea (ENU) have also been used effectively (Gilchrist and Haughn, 2005). They typically introduce small alkyl groups onto the bases themselves, thus modifying their capacity for base pairing. An alkylated guanine (G) pairs with thymine (T), rather than with cytosine (C) (Ashburner, 1990). Thus, if this lesion is not repaired before the next replication, a T will be incorporated instead of a C. Following the second replication event, the original GC base pair will be replaced by an AT, resulting in GC → AT transition. Modification of T by some alkylating agents can also produce TA → CG transitions by a similar mechanism. Accordingly, a substantial fraction of the mutants produced by alkylating mutagens are either missense or nonsense mutations (Greene et al., 2003; Henikoff et al., 2003).

The second class of chemical mutagens is the cross-linking agents. Cross-linking mutagens such as diepoxybutane (DEB) and *cis*-platinum (II) diaminodichloride (*cis*-platinum) can chemically interlock two sites on the DNA molecules. Cross-links that connect two sites on the same strand are known as intra-strand cross-links and those that connect opposite strands are referred to as inter-strand cross-links. Both types of cross-links prevent replication and are difficult for the cell to repair. When

incorrectly repaired, such lesions often produce small deletions.

Physical and chemical alkylating mutagens have a different mutant spectrum, which is most obvious when one compares the ratio M₁ sterility versus embryo-lethal and chlorophyll mutants. For instance, in an experiment where EMS and X-rays were compared, an EMS treatment resulting in 44% sterility (percentage reduction of fertilized ovules compared to control) gave 18.2% embryonic lethals and 9.4% chlorophyll mutants, whereas an X-ray treatment that gave almost the same percentage of sterility produced only 5.9% embryo-lethals and 1.0% chlorophyll mutants (Van der Veen, 1966). This relative high sterility is due to the fact that irradiation leads to chromosome breaks and chromosomal aberrations, giving rise to meiotic disturbances and, therefore, sterility. Base pair substitutions caused by alkylating agents may lead to specific amino acid changes, which may alter the function of proteins but do not abolish their functions as deletions and frame-shift-mutations mostly do. Plant species differ in their sensitivity for specific chemicals. In barley, for example, sodium azide is a very potent mutagen, but it is hardly effective in *Arabidopsis*, where EMS is preferred (Gichner and Veleminsky, 1977; Kleinhofs et al., 1978). However, several other chemicals have shown to be effective in *Arabidopsis*.

Mutagen Dose

The higher the mutagen dose, the more efficient the experiment, and the higher the number of possible mutants. However, there are good reasons not to use the highest mutagen dose possible, since this also will lead to an increase in unwanted mutations at other loci, many of which lead to sterility or even lethality. This makes a relatively low dose attractive in those cases where the amount of work and costs involved in growing more M₁

plants and screening more M_2 plants is not excessive. In those cases where it is expensive to grow many M_1 plants and when laborious mutant screens are applied, the costs involved justify a higher dose. This approach then leads to more 'dirty' mutants that need 'cleaning up' by subsequent backcrosses to the WT.

A simple theoretical model (Malmberg, 1993) predicts that the optimum yield of mutations for a given mutagen should be the dose when 37% of the M_1 seeds survive to reproduction and contribute to the M_2 . This may be a useful starting point but the ultimate measure of success is the production of mutants. Two traits for calibrating the effectiveness of mutagen and dose are the frequencies of embryonic lethals and the production of albino seedlings. Both traits are readily scored and should be abundant, as they represent multilocus targets. Embryonic lethals can be scored directly in the siliques of the M_1 plants, appearing as a 3:1 segregation of pale white embryos instead of the normal green. After EMS mutagenesis, the percentage of M_1 plants segregating embryonic lethals can be as high as 5-10%.

Mutation Frequencies

Mutation frequencies (m) should be based on the frequency at which mutations occur per treated cell. Usually, these are expressed per locus or per group of mutants. In practice, these estimates are based on the estimate of the number of mutants found in the M_2 in relation to the number of M_2 plants that are screened. It is also possible to base the estimate on the number of M_1 plants whose progeny were screened for mutants.

The method of Gaul (Koornneef et al., 1982; Yonezawa and Yamagata, 1975) is the most simple and expresses the number of mutants found per number of M_2 plants divided by f , the average mutant frequency in the progeny of a

heterozygote ($f = 0.25$ for recessive mutations). Therefore, $m = m' / n \times f$ in which m' is the number of mutants found per n M_2 plants screened. This estimate is independent of the number of M_1 plants from which these are derived and also from the degree of chimerism. By dividing m by two the frequency of mutation per haploid genome can be obtained.

Li and Rédei (1969) derived the same frequency from the number of M_1 plants whose progeny were tested. This procedure takes into account the GECN and also requires that all mutations will be detected and that, therefore, the size of individual M_1 progenies should be large enough not to miss the mutant by chance. When $GECN = 2$, the mutant occurs only at a frequency of 1/8 which requires 23 plants not to miss the aa plant by chance in a given progeny at $P < 0.05$. According to the method of Li and Rédei the mutation frequency per cell is calculated with the formula: $m = M / (S \times GECN)$, where M is the number of M_1 progeny segregating for a mutation (type) among S progeny that were screened.

Haughn and Sommerville (1987) have estimated that an M_1 population size of roughly 125,000 plants, after EMS mutagenesis, will effectively saturate the genome for all possible EMS-inducible base changes. Rédei (1975) describes similar calculations to optimize mutant isolation strategies.

Using various EMS-based protocols, mutation frequencies ranging from 1/300 to 1/30,000 per locus have been reported. On the basis of these mutation rates, a GECN value of 2 and the Poisson distribution (Rédei and Koncz, 1992), the number of M_1 plants required is ~700 and ~70,000 respectively.

In practice, many screens are based on 2,000-3,000 M_1 plants after exposure to 15-20 mM EMS for 12 hours, which is at sub-saturation level but typically yields three or more alleles per locus for easily scorable floral mutants (Bowman, non published results).

INSERTIONAL MUTAGENESIS

Although a detailed explanation of insertional mutagenesis goes beyond the scope of this thesis, a brief mention, referring to some excellent reviews on this subject, is warranted.

Currently, T-DNA and transposons are the two main insertional mutagens widely used for gene disruption in *Arabidopsis* (Fedoroff, 1989; Errampalli et al., 1991; Bancroft and Dean, 1993a; Fedoroff and Smith, 1993; Apiroz-Leehan and Feldmann, 1997; Martienssen, 1998; Krysan et al., 1999; Parinov and Sundaresan, 2000).

An advantage of using T-DNA as an insertional mutagen is that it directly generates stable insertions into genomic DNA and does not require additional steps to stabilize the insert. On the other hand complex patterns of T-DNA integration, including transfer of vector sequences adjacent to T-DNA borders and the large frequency of concatemeric T-DNA insertions, can complicate further PCR analysis (De Buck et al., 2000; Kononov et al., 1997). Small and major chromosomal rearrangements induced by T-DNA integration have been frequently observed, leading to difficulties in the genetic analysis of the insertion, such as mutant phenotypes that are not correlated with the T-DNA insertion (Nacry et al., 1998; Laufs et al., 1999). Nevertheless, the development of highly efficient methods of T-DNA transformation for *Arabidopsis* has made it feasible for laboratories to generate thousands of transformants relatively rapidly (Krysan et al., 1999; Clough et al., 1998).

While T-DNA insertions are easily generated in *Arabidopsis*, it is difficult to generate a large collection of independent T-DNA lines in plant species for which transformation methods are more laborious. In contrast, transposon mutagenesis can be accomplished using a limited number of starter lines generated by transformation (Ramachandran and Sundaresan, 2001). Transposons integrate in the genome as single intact elements,

as compared to the complex integration patterns that are frequently generated by T-DNA, and can be used for reversion analysis by remobilizing the transposon insertion. It is well established that transposons of the Ac-Ds family preferentially jump to linked sites (Bancroft and Dean, 1993b; Machida et al., 1997; Ramachandran and Sundaresan, 2001). This property of transposable elements can be exploited for targeted (i.e. localized) transposon mutagenesis of a chromosomal region, which has been used in endogenous systems (Sundaresan, 1996) as well as in heterologous systems (Smith et al., 1996; Muskett et al., 2003; Zhang et al., 2003). Localized transposon mutagenesis can also be used to generate double knockouts of tandemly duplicated genes through reinsertion of transposed elements (Tantikanjana et al., 2004), which would otherwise require a tedious screen for recombinants of two closely linked insertions within the duplicated genes.

1.3.2.2 THE DESIGN OF FORWARD GENETIC SCREENS

A genetic screen is a procedure to identify and select individuals which possess a phenotype of interest. A screen for new genes based on the observation of a particular mutant phenotype is often referred to as a forward genetic screen. On the other hand, identifying new mutant phenotypes caused by disruption of genes that have already known sequences is called a reverse genetic screen. Although the complete sequence of the *Arabidopsis* genome is known, its functional characterization depends on well-designed forward genetic screens, which remain a powerful strategy to identify genes that are involved in many aspects of the plant life cycle.

The pioneers of *Arabidopsis* genetics carried out screens in strains that were collected from the wild and isolated mutants that showed altered leaf shape, altered flowering time and flower morphology (McKelvie, 1962) or that were

characterized by embryo (Meinke and Sussex, 1979) or seedling lethality (Mayer et al., 1991). In addition, screening on defined media allowed the recovery of mutants that were defective in phytohormone synthesis and perception (Bleecker et al., 1988; Guzman and Ecker, 1990; Koornneef et al., 1982; Koornneef et al., 1984; Maher and Martindale, 1980) and other biosynthetic pathways (Li and Rédei, 1969). Screens that use brute force biochemical assays and screening under specific biotic and abiotic environmental conditions led to the analysis of the lipid biosynthesis pathway (Browse et al., 1985; James and Dooner, 1990; Lemieux et al., 1990) and to the genetic analysis of photomorphogenesis (Quail, 1998) and plant-pathogen interactions (Glazebrook et al., 1997), respectively. These screens, however, only allowed identifying obvious phenotypes. Therefore, genetic screens became more sophisticated through the design of particular genetic backgrounds in which more precise questions can be addressed.

The success of a forward genetic screen depends mainly on the definition of a suitable genetic background and an easy and tight procedure to identify the mutants of interest.

Whereas a WT background is suitable for recovering mutants with striking phenotypes, it is not suitable for detecting additional loci that are redundant to a previously identified gene. Neither does a WT background allow the identification of mutants that are involved in processes that do not immediately lead to a morphological or easily scorable biochemical phenotype. At present, the establishment of transgenic approaches that use drug or herbicide resistance genes for selection, and reporter genes to construct visible phenotypes, allows *Arabidopsis* researchers to engineer a scorable phenotype for virtually any process in the life cycle of a plant.

The simpler the screening procedure, the better suited it will be for screening large numbers of plants, thereby increasing the chances of identifying

mutants of interest. Ideally, if a selection method is available, this is preferable to screening, as it is less laborious. If screening cannot be avoided, it should be easy and tight (the fraction of false positives should be low). If neither selection nor an easy screening procedure is possible, the assay should at least be very informative, as is the case for enhancer detection and gene trapping, which provide detailed information about the expression of a gene. Two-step screens offer a successful way to isolate mutants that are not easy to identify (Bartee et al., 2001; Moore et al., 1997). The first step consists of screening both for an easily scorable trait that can be expected to alter in the mutant candidates and for a secondary phenotype that is characteristic for the mutants targeted. As the secondary phenotype could also result from phenomena that are unrelated to the biological process of interest, the mutant candidates must be further characterized in a second step using the more laborious assay to select the informative mutants.

1.3.2.3 THE ISOLATION OF TAGGED MUTATED GENES

Using T-DNA tagging, several *Arabidopsis* genes have been identified and cloned. Once co-segregation of T-DNA along with the phenotype has been established for a given mutant, isolation of the disrupted gene can be achieved by a number of strategies.

Sequences flanking the insertion can be easily identified from single or low copy number lines using inverse PCR (iPCR) and thermal asymmetric interlaced (TAIL) PCR (Liu et al. 1995). Recently, Antal et al. (2004) have reported a simplified version of TAIL PCR, called single oligonucleotide nested (SON)-PCR, which involves only two rounds of PCR with two or three nested sequence primers.

Another strategy employed, called plasmid rescue (Yanofsky et al. 1990; Feldmann, 1992), makes use of insertion within T-DNA of an 'ori' sequence, a

replication of origin required for replication of plasmids in *E. coli*, along with a bacterial selection marker.

Finally, T-linker specific ligation PCR (T-linker PCR), described by Yuanxin et al. (2003), provides another simple and efficient way to amplify the T-DNA flanking region.

1.3.2.4 THE ISOLATION OF UNTAGGED MUTATED GENES USING MAP-BASED CLONING APPROACHES

If a mutation is caused by the insertion of a T-DNA or transposon, the inserted sequence provides a tag pointing directly to the gene. However, when radiation or chemical agents are used a map-based cloning (or positional cloning) approach is required for the isolation of the mutated genes. Although this procedure is more laborious and time-consuming, there are several good reasons to keep screening for radiation or chemically induced mutants.

Besides from being easy to apply and inducing a high number of mutations per plant, the use of radiation or chemical mutagens has additional advantages. For instance, mutations with special properties such as weak, dominant or conditional alleles are more likely to be created using radiation or chemical agents as mutagens. This is in contrast to insertional mutagenesis, which tends to result in complete knockouts of a gene, making it difficult to associate a phenotype other than death with essential genes. It was reported that leaky mutations in *VTC1* (*CYT1*) can result in ozone sensitivity and reduced vitamin C levels in *Arabidopsis*, while knockout mutations cause embryo lethality (Conklin et al., 1999; Lukowitz et al., 2001). Moreover, key regulatory steps in biochemical pathways are often found through dominant point mutations that prevent feedback inhibition of an enzyme, such as anthranilate synthase (Kreps et al., 1996; Li and Last, 1996) or aspartic acid kinase (Heremans and Jacobs, 1997). Such dominant

mutations would not be found by using insertional mutagenesis. Also, many chemically induced second-site mutations have been identified.

The big advantage of map-based cloning is that it is an approach without prior assumptions. It is a process of discovery that makes it possible to find mutations anywhere in the genome, including intergenic regions and genes that do not resemble any gene with known or inferred function. A prerequisite for map-based cloning is the availability of a genome map. Such a map shows the positions of genes and other sequence features, which serve as reference points and guide one to the region containing the gene of interest. Conventionally, the construction of genome maps, or genome mapping, has been divided into two categories: genetic mapping and physical mapping.

GENETIC MAPPING

Genetic mapping uses genetic techniques, such as cross-breeding and pedigree analysis, to construct maps showing the positions of genes and other sequence features on a genome. The principles of inheritance as first described by Mendel in 1865 were the first step towards genetic mapping.

Mendel's Laws of Inheritance

From the results of his breeding experiments with peas, Mendel concluded that each pea plant possesses two alleles for each gene, but displays only one phenotype. This is easy to understand if the plant is homozygous for a particular trait, as then it possesses two identical alleles and displays the appropriate phenotype. However, Mendel showed that if two pure-breeding plants with different phenotypes are crossed, all the progeny (F_1 generation) display the same phenotype.

These F₁ plants were heterozygous for the phenotypes being studied, since each inherited one allele from each parent. The phenotype expressed in the F₁ plants is dominant over the second, recessive phenotype.

Mendel carried out additional crosses that resulted in the postulation of two laws of genetics. The First Mendelian Law states that alleles segregate randomly (figure 1.12). This means that each member of an F₁ progeny has the same chance of inheriting one parent's allele as it has of inheriting the other parent's allele. The Second Law described by Mendel is that pairs of alleles segregate independently, so that the inheritance of alleles of one gene is independent of the inheritance of alleles of another gene. Mendel's theory about the existence of 'unit factors' (genes) and the way they are inherited was absolutely correct.

We know now that the two alleles of a gene are carried by the two homologous chromosomes in a diploid cell and that the way in which alleles are transmitted from parent to offspring corresponds to the events that occur when haploid gametes are produced during meiosis. The simple dominant-recessive rule proposed by Mendel, however, can be complicated by situations he did not encounter with the pea plants. One of these is incomplete dominance, where the heterozygous phenotype is intermediate between the two homozygous forms. Another complication is codominance, when both alleles are detectable in the heterozygote. Mendel made only one major error: his Second Law does not allow for linkage, the possibility that alleles of two genes will be inherited together because they are on the same chromosome. The discovery of linkage by Mendel's followers led to the methods used for genetic mapping.

The Discovery of Linkage

As early as 1903, Walter S. Sutton, who along with Theodor Boveri united the

fields of cytology and genetics, pointed out the likelihood that in organisms there are many more 'unit factors' (later called alleles) than chromosomes. Soon thereafter, genetic investigations with several organisms revealed that certain genes were not transmitted according to Mendel's second law of independent assortment. When studied together in matings, these genes seemed to segregate as if they were somehow joined or linked together. Further research showed that such genes were part of the same chromosome and they were indeed transmitted as a single unit. We know now that most chromosomes consist of very large numbers of genes and, in fact, contain sufficient DNA to encode thousands of these units. Genes that are part of the same chromosome were said to be linked and to demonstrate linkage in genetic crosses. However, when the early geneticists carried out genetic crosses similar to those published 40 years earlier by Mendel, they discovered that very few pairs of genes displayed complete linkage. Pairs of genes were either inherited independently, as expected for genes in different chromosomes, or if they showed linkage then it was only partial linkage: sometimes they were inherited together and sometimes they were not (figure 1.13). The resolution of this contradiction between theory and observation was the critical step in the development of genetic mapping techniques.

The Behavior of Chromosomes during Meiosis Explains Why Genes Display Partial and Not Complete Linkage

The *Drosophila* geneticist, Thomas H. Morgan, first discovered the phenomenon of X-linkage. In his studies, he investigated numerous *Drosophila* mutations located on the X chromosome. When he analyzed crosses involving only one trait, he was able to deduce the

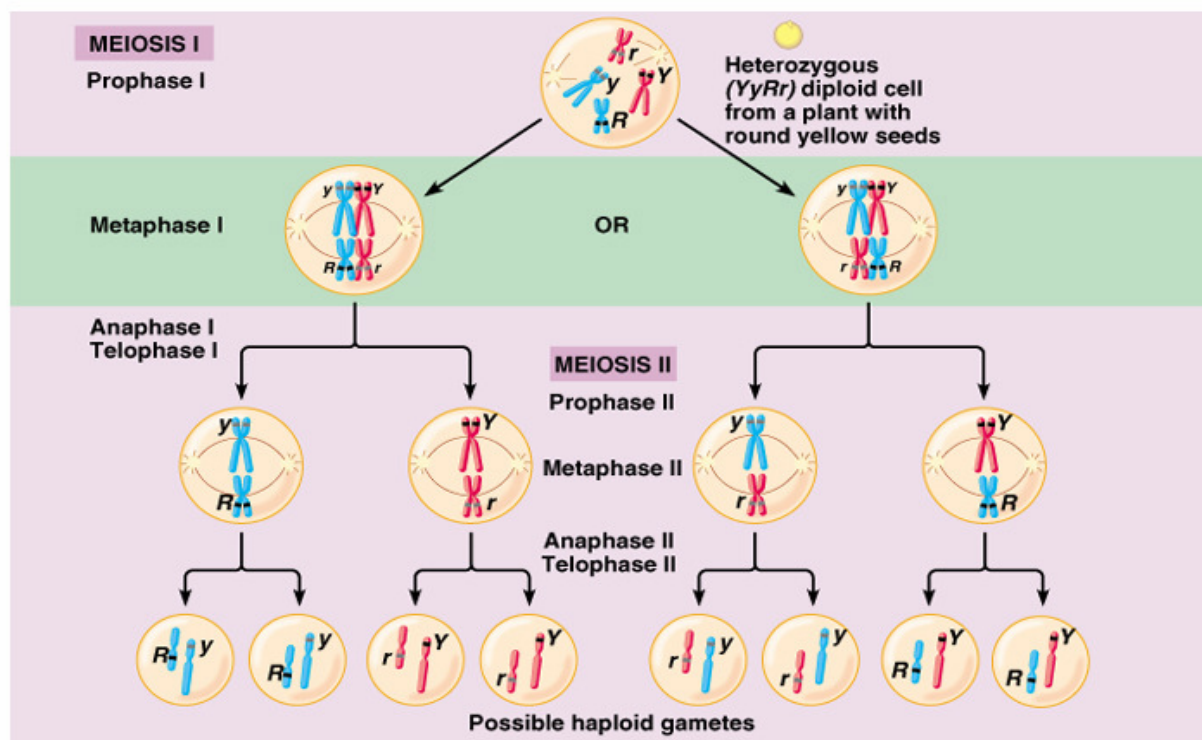


FIGURE 1.12 THE MEIOTIC BASIS FOR MENDEL'S LAW OF INDEPENDENT ASSORTMENT (GRIFFITHS ET AL., 1996). Independent assortment of the alleles of two genes on different chromosomes is illustrated by meiosis in a pea plant heterozygous for seed color (Yy) and seed shape (Rr). Pea plants have seven pairs of chromosomes, but only the two pairs bearing the seed color alleles and the seed shape alleles are shown. During meiosis, segregation of the seed color alleles occurs independently of segregation of the seed shape alleles. The basis of this independent assortment is found at metaphase I, when the homologous pairs (bivalents) align at the spindle equator. Each bivalent can face in either direction. Consequently, there are two alternative situations: either the paternal homologues (blue) can both face the same pole of the cell, with the maternal homologues (red) facing the other pole, or they can face different poles. The arrangement at metaphase I determines which homologues subsequently go to which daughter cells. Since the alternative arrangements occur with equal probability, all the possible combinations of the alleles therefore occur with equal probability in the gametes. For simplicity, chiasmata and crossing over are not shown.

mode of X-linked inheritance. However, when he made crosses that simultaneously involved two X-linked genes, his results were at first puzzling.

In 1911, Morgan proposed that the gene separation he observed could be explained based on the cytological studies of F. Janssens. Janssens had observed that synapsed homologous chromosomes in meiosis wrapped around each other, creating chiasmata where points of overlap are evident. Morgan stated that these chiasmata could represent points of genetic exchange. After performing numerous crossing experiments, he concluded that linked genes exist in a linear order along the chromosome and that a variable amount

of exchange occurs between any two genes. Furthermore, he explained that the closer two genes are on a chromosome, the less likely it is that a chiasma is formed between them and that a genetic exchange will occur. The term crossing over was first proposed by Morgan to describe the physical exchange leading to recombination (Griffiths et al., 1996).

Nowadays, crossing over is viewed as an actual physical breaking and rejoining process that occurs during the meiotic prophase. The double-strand-break-repair model (Szostak et al., 1983) is acknowledged to best explain meiotic reciprocal recombination. In this model two sister chromatids break at the same point and their 5' ends are degraded. In the next step one of the 3' tails invades

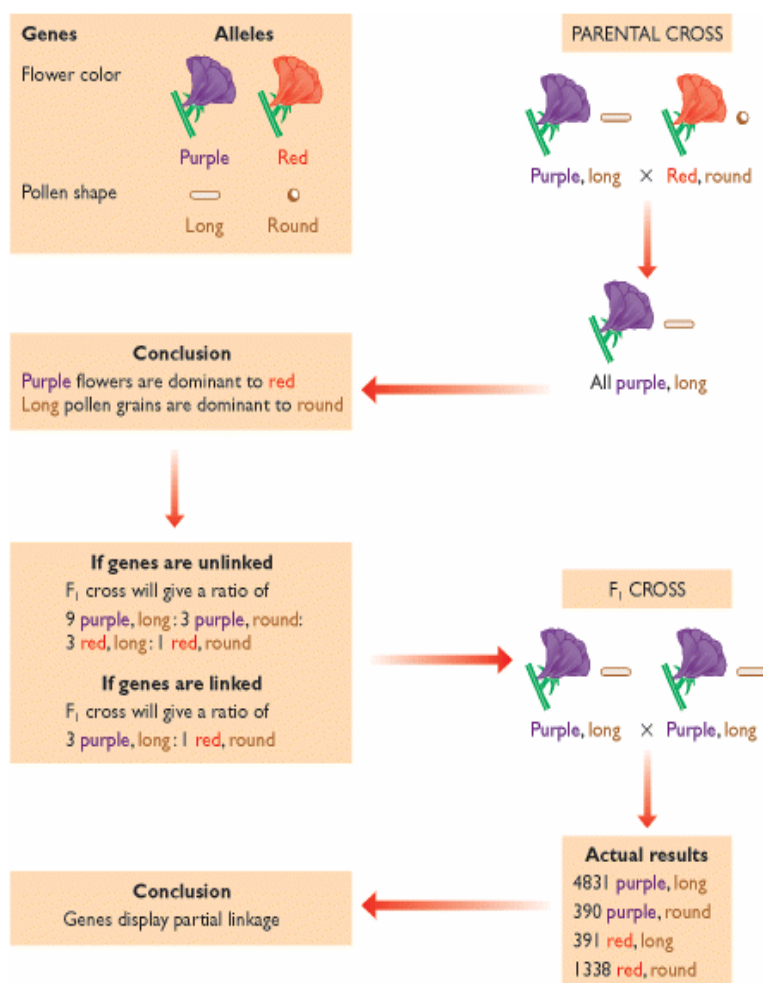


FIGURE 1.13 PARTIAL LINKAGE (GRIFFITHS ET AL., 1996). Partial linkage was discovered in the early 20th century. The cross shown here was carried out by Bateson, Saunders and Punnett in 1905 with sweet peas. The parental cross gives the typical dihybrid result, with all the F₁ plants displaying the same phenotype, indicating that the dominant alleles are purple flowers and long pollen grains. The F₁ cross gives unexpected results as the progeny show neither a 9 : 3 : 3 : 1 ratio (expected for genes on different chromosomes) nor a 3 : 1 ratio (expected if the genes are completely linked). An unusual ratio is typical of partial linkage.

the intact homolog, displacing one of the strands. This 3' end serves as a primer for polymerase, which extends the tail until it can eventually be joined to a 5' end. Meanwhile, the displaced strand serves as a template to fill the gap left in the other broken strand. Two Holliday junctions form and may produce recombinant flanking DNA, depending on how they are resolved.

Crossing-over provides an enormous potential for genetic variation in the gametes formed by any individual. This type of variation, together with that

resulting from independent assortment of the chromosomes, ensures that all offspring will contain a diverse mixture of maternal and paternal alleles.

From Partial Linkage to Genetic Mapping

Morgan's student, Alfred H. Sturtevant, was the first to realize that his mentor's proposal could also be used to map the sequence of linked genes. Sturtevant knew from Morgan's work that the frequency of exchange could be

taken as an estimate of the relative distance between two genes. He argued that the recombination frequencies between linked genes are additive and constructed a map of three X-linked *Drosophila* genes. He defined one genetic map unit as that distance between genes for which one product of meiosis out of hundred is recombinant. So, one map unit corresponds to a recombination frequency of 0.01 (or one percent) and is now referred to as one centiMorgan (cM) in honor of Morgan's work. In addition to these three genes, Sturtevant considered two other genes on the X-chromosome and produced a more extensive map including all five genes. He and a colleague, Calvin Bridges, soon began a search for autosomal linkage in *Drosophila*. By 1923, they had clearly shown that linkage and crossing over were not restricted to X-linked genes. Rather, they had discovered linked genes on autosomes, between which crossing over occurred (Griffiths et al., 1996).

Although Sturtevant's basic principles are accepted as correct, it turns out that his assumption about the randomness of crossovers was not entirely justified. Comparisons between genetic maps and the actual positions of genes on DNA molecules, as revealed by physical mapping and DNA sequencing, have shown that some regions of chromosomes, called recombination hot spots, are more likely to be involved in crossovers than others. This means that a genetic map distance does not necessarily indicate the physical distance between two genes or DNA features. Also, we now realize that a single chromatid can participate in more than one crossover at the same time, but that there are limitations on how close together these crossovers can be, leading to more inaccuracies in the mapping procedure. Despite these qualifications, linkage analysis usually makes correct deductions about gene order, and distance estimates are sufficiently accurate to generate genetic maps that are of value as frameworks for genome sequencing projects.

Markers For Genetic Maps

The First Genetic Maps Were Constructed Using Genes as Markers

The first genetic maps, constructed in the early part of the 20th century for organisms such as *Drosophila*, used genes as markers. However, geneticists soon realized that only a limited number of genes were responsible for visual phenotypes whose inheritance can be studied, and in many cases the analysis is made less easy because more than one gene affects a single feature. To make genetic maps more comprehensive, it became essential to find characteristics that were more numerous, more distinctive and less complex than visual ones. Especially in bacteria and yeast, where few visual characteristics are present, using biochemistry to distinguish phenotypes was a solution to this problem. Also in humans biochemical phenotypes scorable by blood typing have been studied in the 1920's (Yamamoto et al., 1990). Although genes are very useful markers, they are by no means ideal. Especially with larger genomes such as those of vertebrates and flowering plants, a map based entirely on genes is not very detailed.

Molecular Markers for Genetic Mapping

Restriction fragment length polymorphism (RFLP) markers have been used extensively for the construction of genetic maps (table 1.1). RFLPs are variations in the length of the DNA fragments produced after cleavage with a restriction enzyme (Tanksley et al., 1989). Among the various molecular markers developed, they were the first to be used in human genome mapping (Botstein et al., 1980) and later they were adopted for plant genome mapping (Weber and Helentjaris, 1989). Rice is the most

extensively analyzed species by RFLPs (Causse et al., 1994; Kurata et al., 1994).

The first technique used to type RFLPs was Southern hybridization. The DNA fragments resulting from restriction are separated from one another by agarose gel electrophoresis and the ones being studied detected by hybridization probing. The polymorphism detected by RFLP is very reliable as it involves the recognition by specific restriction enzymes and hybridization with a specific probe. RFLPs are codominant since both alleles can be directly detected. However, such RFLP analysis is labour intensive and time consuming.

PCR is now more commonly used to detect RFLPs (Konieczny and Ausubel, 1993). Here, the region of DNA containing the RFLP is amplified before treatment with the restriction enzyme. The amplified fragments, called cleaved amplified polymorphic sequences (CAPS) can then be visualized on agarose gel using ethidium bromide. This is a less sensitive detection method than hybridization probing so that the background of unamplified fragments does not show up.

Another well known probe-based marker is the variable number tandem repeat (VNTR) or minisatellite, first discovered in humans by Alec Jeffreys in the 1980's, and extensively used in forensic DNA profiling (Jeffreys et al., 1985). VNTRs consist of tandem arrays of short repeated sequences (a few tens of nucleotides in length) highly dispersed throughout the genome at numerous loci. They are embedded in unique flanking sequences and the loci are hypervariable in terms of their number of repeat units. Not only are there many different loci, but multi-allelic forms of single loci exist as well at the population level due to unequal crossing over.

VNTRs tend to be found more frequently near the ends of chromosomes, which is less convenient. Most of the VNTRs are longer than 300 base pairs since the repeat units are relatively long and there tend to be many of them in a single array.

Therefore they are less amenable to PCR and are detected by hybridization probing.

Random amplified polymorphic DNA (RAPD) analysis involves amplification with PCR using a single or a combination of typically 10-mer oligonucleotides (Welsh and McClelland, 1990; Williams et al., 1990). Variation is based on the position and orientation of primer-annealing sites and the interval they span.

The amplification products are separated on agarose gel and stained with ethidium bromide. RAPDs are dominant in the sense that the presence of a RAPD band does not allow distinction between hetero- and homozygous states. Their main disadvantages are poor reliability and reproducibility and their sensitivity to experimental conditions (Karp et al., 1996).

Simple sequence length polymorphisms (SSLPs) (also termed simple sequence repeats (SSRs), short tandem repeats (STRs), sequence tagged microsatellite sites (STMS) or microsatellites) are a type of insertion/deletion polymorphisms (InDels) and exploit the variability of short repetitive sequences for mapping purposes (Bell and Ecker, 1994). A primer pair is used to amplify a fragment containing a short repetitive element. The length of this repeat differs between different alleles. Consequently, PCR products will also differ in length. Typically, SSLPs may be dinucleotides ((AC)_n, (AG)_n, (AT)_n), trinucleotides ((TCT)_n, (TTG)_n), tetranucleotides ((TATG)_n) and so on, where n is the number of repeating units of the microsatellite. In addition to occurring at many different loci, they can also be polyallelic. SSLPs have been applied extensively towards genetic and physical mapping in plant and animal species. (AT)_n dinucleotides are the most abundant type of SSLP in plants. They are co-dominant and their detection on agarose gel is inexpensive and straightforward. However, specific primers need to be designed and tested, which is a time-consuming and costly business in species

of which the genome sequence is not known. The use of random primers, as in RAPD analysis, avoids this up-front input. Also non-repetitive InDels are used as molecular markers for genetic mapping. For instance for *Arabidopsis*, 18579 Indels have been identified (<http://www.arabidopsis.org/Cereon/index.html>), of which only 10% were associated with simple sequence repeats but most were found in non-repetitive sequences.

Sequenced characterized amplified regions (SCARs) are DNA fragments amplified by PCR using specific 15-30 base pair primers, designed from nucleotide sequences established in cloned RAPD fragments. By using longer PCR primers, SCARs do not face the problem of low reproducibility generally encountered with RAPDs. Obtaining a codominant marker may be an additional advantage of converting RAPDs into SCARs. However, SCARs may exhibit dominance when one or both primers partially overlap the site of sequence variation. SCARs are locus specific and have been widely applied in gene mapping studies (Nair et al., 1995; Nair et al., 1996; Paran and Michelmore, 1993).

The amplified fragment length polymorphism method (AFLP) combines the use of restriction enzymes with PCR amplification of fragments and detects either a single nucleotide polymorphism (SNP) in or immediately adjacent to one or both restriction sites, or an insertion/deletion (InDel) event in the sequence of one of the alleles (table 1.1; Vos et al., 1995). The AFLP technique is highly reproducible and requires no prior sequence information. It can thus be applied directly to any organism. Genomic restriction fragments are adaptor-ligated and PCR primers are designed based on adaptor and restriction site sequences. Subsets of fragments can subsequently be specifically amplified by adding random selective nucleotides at the 3' end of the

primers. The number of selective nucleotides needed depends on the size of the genome studied. Thus, although it is based on restriction and random selection, the targeted subsets of PCR templates are highly specific and the procedure is robust. A great advantage is that one basic set of primers can be used for organisms with comparable genome sizes.

Single nucleotide polymorphisms (SNPs) are individual point mutations in the genome (table 1.1). There are vast numbers of these, some of which give rise to RFLPs or CAPS markers, but many of which do not because the sequence in which they lie is not recognized by any restriction enzyme. Derived cleaved amplified polymorphic sequences (dCAPS) (Michaels and Amasino, 1998; Neff et al., 1998) can exploit almost all single nucleotide changes for mapping purposes. To achieve this, a mismatched PCR primer is designed next to the polymorphic position such that an artificial restriction site is created. Consequently, at this position the sequence of all PCR products is changed with respect to the genomic sequence. Restriction sites with interrupted palindromes are frequently used to design dCAPS markers because they allow to position the mismatched nucleotide at a distance from the 3' end of the oligonucleotide where it is less likely to interfere with the priming of *Taq* polymerase. The size difference after digestion is rather small and essentially determined by the length of the mismatched primer. Disadvantages of using CAPS and dCAPS for genotyping include the extra time and cost involved in the restriction enzyme digestion and the possibility of a false result attributable to incomplete digestion by the restriction enzyme.

It is also possible to detect SNPs using allele-specific PCR primers, where the 3' end of a primer has a perfect match with one allele and a mismatch with the other allele (Ugozzoli and Wallace, 1991).

TABLE 1.1 THE MAIN CATEGORIES OF MOLECULAR MARKERS.

MARKER NAME	VARIATION TYPE			INFORMATION CONTENT		
	SNP ¹	INDEL ²	VNTR ³	2 DOMINANT ALLELES	2 CO-DOMINANT ALLELES	MULTI-ALLELIC CO-DOMINANT
RFLP	+	(+) ⁴	(+)	-	+	(+) ⁵
PCR-RFLP	+	(+) ⁴	(+)	-	+	-
RAPD	+	(+) ⁴	(+)	+	-	-
AFLP	+	(+) ⁴	(+)	+	(+) ⁶	-
SSCP ⁹	+	(+) ⁴	(+)	-	+	(+) ⁵
Microsatellite	-	(+) ⁷	+	-	-	+
SNP	+	(+) ⁸	-	-	+	-

¹ Single Nucleotide Polymorphism: any kind of base substitution. The fact that SNPs appear both as a variation type and a marker name is due to the fact that in reality, many techniques used for genotyping SNPs are grouped under this generic marker name.

² Insertions and deletions.

³ Variable number of tandem repeats.

⁴ Although the RAPD, AFLP, RFLP, PCR-RFLP and SSCP techniques will detect base substitutions in the vast majority of cases, the two other types of DNA variation can also be analyzed.

⁵ In some instances, more than two alleles can be analyzed.

⁶ With an automatic sequencer, some markers can be scored as co-dominant.

⁷ Variations in PCR product length can be due to a deletion in the sequence flanking the microsatellite.

⁸ Many SNP detection techniques can also be used for scoring small insertions or deletions (indels).

⁹ Single strand conformation polymorphism (SSCP) is based on differences in folding conformation of single stranded DNA, when placed in non-denaturing conditions.

In theory, such primers can be used to preferentially amplify one allele of a SNP, but in practice a single-basepair change is often not enough to allow reliable differentiation between the two alleles of a SNP (Kwok et al., 1990; Cha et al., 1992). Drenkard et al. described a modification of the allele-specific amplification procedure (single nucleotide amplified polymorphism (SNAP)) (Drenkard et al., 2000). In this method, additional mismatches are introduced in the amplifying primers to maximize the difference in the amplification efficiencies of the two alleles of the SNP. Primer base pair changes that allow differential amplification of SNP sites can be predicted using the SNAPER program (<http://patho.mgh.harvard.edu/ausubelweb>). As with CAPS, SNAP markers are codominant and can be detected on agarose gels. However, it is necessary to run two PCR reactions (one for each allele

of the SNP) to get complete SNAP genotyping data.

SNPs can also be typed by methods that do not involve gel electrophoresis. This is important because gel electrophoresis has proved to be difficult to automate so any detection method that uses it will be relatively slow and labor intensive.

Oligonucleotide arrays (Gene Chips) contain thousands of oligonucleotides annealed to a glass slide. Such arrays allow the detection of SNP polymorphisms by differential hybridization in a highly parallel and automated manner (Lipshutz et al., 1999).

The Taq-Man PCR assay is designed to detect SNPs in a high-throughput manner through the release of fluorescent reporter dye from a quencher on the same oligonucleotide by 5' nuclease activity (Livak, 1999). By using more than one reporter dye, it is possible

to detect different alleles of a SNP in a single reaction. The relatively high price of oligonucleotides tagged with reporter and quencher dyes makes this method cost-effective only if a large number of reactions need to be run with each SNP marker.

In pyrosequencing, an enzymatic cascade and luminometric detection system is used to measure the pyrophosphate that is released as a result of nucleotide incorporation (Ahmadian et al., 2000; Alderborn et al., 2000). Because 20 or more nucleotides are determined by this method, it is possible to detect several closely linked SNPs at once. The pyrosequencing method can be automated but has the disadvantage that it does not work well on stretches of repeated nucleotides.

DHPLC allows the detection of SNPs through different retention time of heteroduplex and homoduplex DNA in reversed-phase HPLC under partially denaturing conditions (Spiegelman et al., 2000). DHPLC allows detection of SNP polymorphisms in PCR-amplified DNA up to about 1,000 base pairs in size.

The method of fluorescence resonance energy transfer combines PCR and oligonucleotide ligation to detect SNPs (Chen et al., 1998). Dye-labeled oligonucleotide probes are used in this assay, and allele-specific ligation is detected by fluorescence resonance energy transfer, which only occurs when two dye-labeled oligos are joined by ligation.

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry can be used to rapidly detect SNPs in short DNA pieces by differences in molecular mass (Wada and Yamamoto, 1997).

A disadvantage of most high-throughput methods for detecting DNA polymorphisms is the high initial equipment cost, which results in a high per-assay cost for a lab that does not need to perform large numbers of genotyping reactions on a routine basis.

PHYSICAL MAPPING

Although genetic mapping has been pursued in plants and animals for decades, it is only relatively recently that advances in cloning and clone fingerprinting have allowed the construction of physical maps (figure 1.14). A physical map is an ordered set of DNA fragments, among which the distances are expressed in physical distance units (base pairs). These days, a physical map usually comprises a set of ordered large insert clones such as bacterial artificial chromosomes (BACs), which have largely replaced yeast artificial chromosomes (YACs) as the preferred building blocks of a physical map. Physical maps can be independent of genetic information but are more valuable if linked to genetically mapped markers, and are even more powerful if integrated with genomic sequence data. A plethora of physical mapping techniques have been developed, including restriction mapping (Meyers et al., 2004; Strachan and Read, 1999), fluorescent *in situ* hybridization (FISH) (Dear, 1997; Heiskanen et al., 1996) and sequence tagged site (STS) mapping (Dear, 1997).

THE ARABIDOPSIS GENOME SEQUENCE

The sequence of the first plant genome was completed in December 2000, and it was the third complete genome of a higher eukaryote (The *Arabidopsis* Genome Initiative, 2000), after *Drosophila melanogaster* (Adams et al., 2000; Myers et al., 2000) and *Caenorhabditis elegans* (C_elegans Sequencing Consortium, 1998).

The *Arabidopsis* sequence represented 115 million base pairs (Mb) of euchromatin out of the estimated 125 Mb total. It was completed by the traditional bacterial artificial chromosome (BAC)-based approach from a minimal tiling path of overlapping large insert clones.

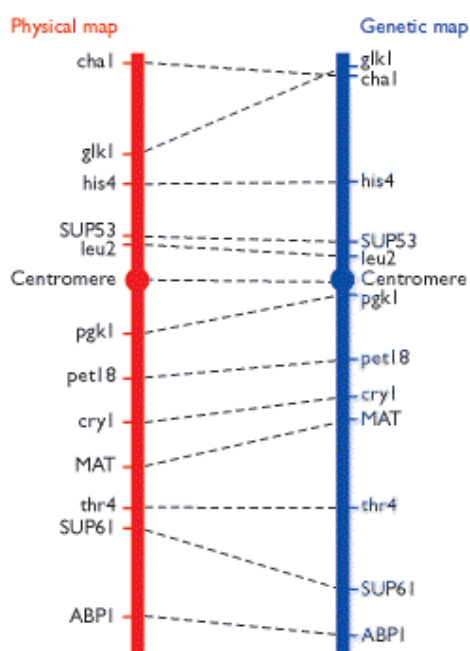


FIGURE 1.14 COMPARISON BETWEEN THE GENETIC AND PHYSICAL MAPS OF *SACCHAROMYCES CEREVISIAE* CHROMOSOME III (OLIVER ET AL., 1992). The comparison shows the discrepancies between the genetic and physical maps, the latter determined by DNA sequencing. Note that the order of the upper two markers (*glk1* and *cha1*) is incorrect on the genetic map, and that there are also differences in the relative positioning of other pairs of markers.

The sequence covered all ten chromosome arms, including parts of the centromeres (figure 3.15). A major finding revealed by the five chromosome sequences was evidence for genome-wide duplication followed by gene loss and major translocations. Tandem duplications are also extensive. Overall, only one of three genes does not have a close family member. The *Arabidopsis* gene repertoire of 11,000–15,000 gene families is comparable in number to other sequenced organisms, highlighting the similarity of life's instructions that stem from our common single-celled ancestors. The total number of *Arabidopsis* genes was initially estimated at 25,490 and later revised to 30,700. What seemed remarkable when the human genome was published (Lander et al., 2001; Venter et al., 2001), was that *Arabidopsis*, the simplest plant genome, had a similar

number of genes to humans. Certainly the number of genes does not reflect the complexity of the organism. Plants often contain many more genes than animals (mainly due to polyploidy or large-scale duplication). More likely, organism complexity is related to the levels of molecular interactions and regulatory circuitry using a similar genetic parts list.

The *Arabidopsis* genome contains an amazing array of genes encoding enzymes involved in primary and secondary metabolism, which is a unique feature of plants. These enzymes are the equipment of the plant chemical factories to build all required metabolic molecules and to generate an arsenal of specialized compounds. Because plants are sessile, they cannot move to avoid biotic attack or abiotic stress, or to find mating partners. Thus, they depend heavily on chemical signals. One example is the large gene family of cytochrome p450s, with more than 300 members involved in small molecule biosynthesis and detoxification (Paquette et al., 2000). *Arabidopsis* also has a large number of transcription factors (approximately 1500), many of which are in families unique to plants, such as the AP2/EREBP, RAV, NAC, ARF, and AUX/IAA families. Plants seem to lack many of the transcription factors families found in animals, such as nuclear steroid receptors. *Arabidopsis* has nearly 1000 serine/threonine kinases and more than 600 are receptor-like ser/thr kinases (Morris and Walker, 2003), but no receptor-like tyrosine kinases were found. Bacterial-like histidine kinases are also present. Some have been recruited as ethylene receptors or as red light photoreceptors, the phytochromes. Although showing sequence homology with histidine kinases, the plant phytochromes and several of the ethylene receptors have evolved ser/thr kinase activity (Xie et al., 2003). Plants have the extraordinary ability to photosynthesize, and hundreds of genes have been identified that are likely involved in light harvesting, chlorophyll biosynthesis, CO₂ fixation, or are a part of the two core energy - generating

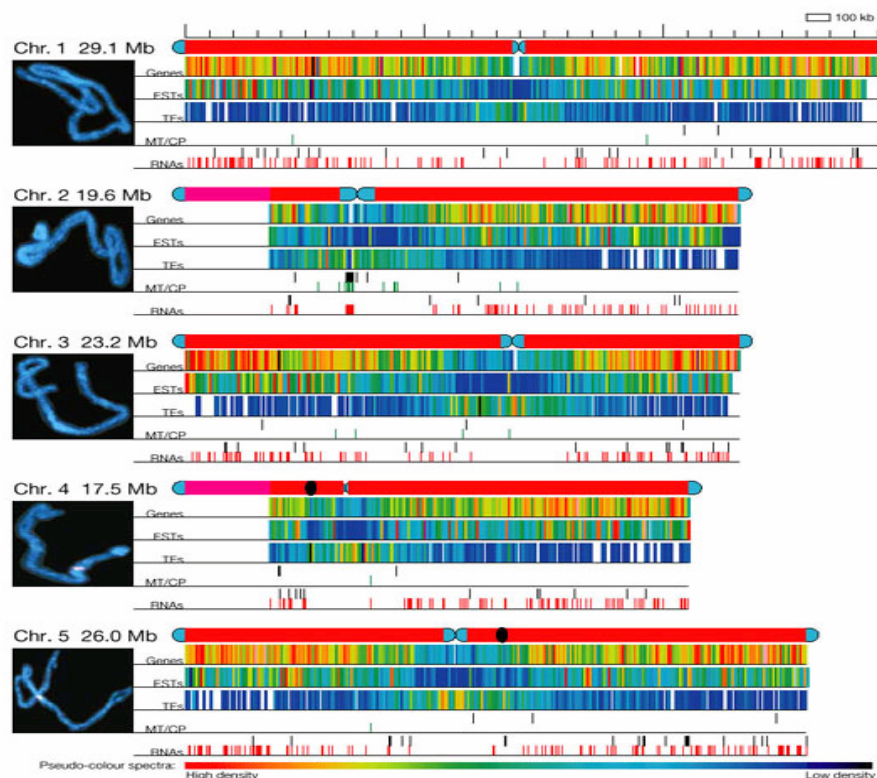


FIGURE 1.15 REPRESENTATION OF THE ARABIDOPSIS CHROMOSOMES (THE ARABIDOPSIS GENOME INITIATIVE, 2000). Each chromosome is depicted as a coloured bar. Sequenced portions are red, telomeric and centromeric regions light blue. Heterochromatic knobs are shown in black and the rDNA repeat regions are magenta. The frequency of features was given pseudo-colour assignments, from red (high density) to deep blue (low density). Gene density ('Genes') ranged from 38 per 100 kb to 1 gene per 100 kb; expressed sequence tag matches ('ESTs') ranged from more than 200 per 100 kb to 1 per 100 kb. Transposable element densities ('TEs') ranged from 33 per 100 kb to 1 per 100 kb. Mitochondrial and chloroplast insertions ('MT/CP') were assigned black and green tick marks, respectively. Transfer RNAs and small nucleolar RNAs ('RNAs') were assigned black and red ticks marks, respectively.

photosystems. Light also regulates development, in a process photomorphogenesis (Neff et al., 2000). The genome contains, in addition to the red light-absorbing phytochromes, blue light-absorbing cryptochrome, and phototropin photoreceptors, as well as hundreds of putative downstream light-signaling proteins (Cashmore, 2003). We know the function of only a few dozen light-signalling proteins. Other unique gene functions of plants and yeast are creating electrochemical gradients using mainly proton-type ATPases, whereas *C. elegans* and *Drosophila* use mainly sodium-type ATPases. Thus, transport is usually coupled with protons rather than sodium ions. In addition, compared to animals, water channels (aquaporins) are highly over-represented in the *Arabidopsis* genome. There are also cellular differences between plants and animals at the genome level. No genes encode intermediate filaments in *Arabidopsis*, whereas actin and α - and β -tubulin are present. The plant cell wall surely is unique, and homologues of animal cytoskeleton

anchorage proteins (which link with the extracellular matrix) have not been seen in the *Arabidopsis* genome. Plants also seem to have a different repertoire of small G-protein-signalling molecules. Different mechanisms are used to protect plants and animals against their biotic environment. No major histoincompatibility complex genes or antibody-like genes were identified in plants; however, plant-specific nucleotide binding site leucine rich repeat (NBS-LRR)-type disease resistance (R) genes are abundant and can be grouped into subfamilies (Meyers et al., 2004). The 149 R genes are present at several loci throughout the *Arabidopsis* genome and occur as singletons or in highly polymorphic clusters. The sequence of the first plant genome revealed extensive horizontal transfer of genes from an ancestral cyanobacterium-like endosymbiont to the plant cell nucleus. New studies show that up to 18% of all plant genes originated from this engulfment, which resulted in the chloroplast (Martin et al., 2002).

Mitochondrial genes have also been transferred to the nucleus. Strikingly, a large translocation of 620 kb brought several duplicated and rearranged copies of the mitochondrial genome near the centromere on chromosome 2 (Stupar et al., 2001). Compared to *Drosophila*, *C. elegans*, and yeast, there are about 150 protein families that are unique to *Arabidopsis*.

MAP-BASED CLONING: GENERAL CONSIDERATIONS

Map-based cloning is the process of identifying the genetic basis of a mutant phenotype by looking for linkage of the mutant phenotype to markers whose physical location in the genome is known. In the past, the paradigm used for map-based cloning has been one in which 'chromosome walking' played a central, and very time-consuming, role. First, either random molecular markers or established molecular linkage maps were used to conduct a search to localize the gene adjacent to one or more markers. A 'chromosome walk' was then initiated from the closest linked marker and a series of overlapping clones (cosmids or YACs) isolated. This continued either until one arrived at another molecular marker known to be situated on the opposite side of the target gene or until there was another indication that the walk had actually passed over the target gene. This was followed by cloning, complementation by transformation and *de novo* determination of the sequence of the entire region of interest to high quality without a previously determined WT DNA sequence as a guide (Arondel et al., 1992; Giraudat et al., 1992; Leung et al., 1994; Meyer et al., 1994; Mindrinos et al., 1994). In large genomes 'chromosome walking' was hampered not only by the large amounts of DNA being traversed, but also by the high frequency of repetitive DNA. Furthermore, libraries of large inserts (YAC libraries) used for 'chromosome walking' in large genomes have often contained

chimeric clones or clones that have become rearranged after cloning.

The strategy of 'chromosome walking' was based on the assumption that it is difficult and time-consuming to find DNA markers that are physically close to a gene of interest. Technological developments invalidate this assumption for many species. As a result, the mapping paradigm has now changed such that one first isolates one or more DNA markers at a physical distance from the targeted gene that is less than the average insert size of the genomic library being used for clone isolation. The markers are then used to screen the library and isolate (or 'land on') the clone containing the gene.

Today, the advances in sequencing projects, the wealth of available marker systems and the progress made in methods to detect DNA polymorphisms make relatively rapid map-based cloning of a gene in model species feasible (Jander et al., 2002).

MAP-BASED CLONING IN ARABIDOPSIS

Map-based cloning of genes in *Arabidopsis* is greatly facilitated by the recent sequencing of the Col-0 and partial sequencing of the Ler genome. These two ecotypes were sequenced because they are among the most commonly used ecotypes in *Arabidopsis* research. Since the work of George Rédei in the 1950's, they have been the subjects of literally thousands of papers that have been published on the genetics, molecular biology, and biochemistry of *Arabidopsis*.

The Col-0 ecotype was the subject of a large international sequencing project, which has produced a nearly complete sequence using a clone-by-clone approach (The *Arabidopsis* Genome Initiative, 2000). Partial genomic sequence data generated from other ecotypes can be positioned on the framework of Col-0 genome sequence and sequencing of individual genes from

mutants or from other ecotypes has become routine. It is simply a matter of designing PCR primers based on the Col-0 sequence, amplifying the desired gene, and sequencing the product.

The sequence of the Ler genome was determined by low coverage shotgun sequencing at Cereon Genomics (Jander et al., 2002). This project generated approximately 700,000 500 base pair sequence traces. More than 200,000 of these traces were chloroplast, mitochondrial, or ribosomal DNA and were not used for the assembly. This left 498,037 traces (or 263 Mbp) of good quality raw sequence. Assembly of the sequences produced 50,262 contigs of an average size of 1.5 kb and 31,044 single-read sequences. The size of the assembled dataset was 92.1 Mbp, suggesting that approximately 70% of the genome is covered at the nucleotide level. It was found that at least some sequence from over 95% of all genes is present in the data assembled from the low coverage shotgun approach.

For *Arabidopsis* researchers who are interested in map-based cloning, the value of two genome sequences greatly exceeds that of only one such sequence. Whereas the availability of the genome sequence of a single ecotype mainly facilitates DNA sequencing in the final stages of a mapping project, data from two genomes make it possible to develop a database of DNA polymorphisms that can be used as genetic markers.

Besides their known genome sequences other available tools and resources make Col and Ler the combination of choice for map-based cloning in *Arabidopsis* (Lukowitz et al., 2000). For the Col accession, for instance, various libraries were made from WT plants and were deposited at the *Arabidopsis* stock centers, including genomic and cDNA phage-libraries, and libraries of yeast- and bacterial artificial chromosomes and of transformation-competent artificial chromosomes, which may be used for molecular

complementation (Liu et al., 1999). Most existing mutations, including mutations causing visible phenotypes that can be employed as genetic markers, have been induced either in a Col or a Ler background. Col and Ler are also the parents of a widely used collection of recombinant inbred lines (Lister and Dean, 1993). Hundreds of markers have been analyzed in these lines, and the genetic map produced from this work has become the standard against which other *Arabidopsis* genetic maps are aligned.

Map-Based Cloning Strategies: Choosing the Appropriate Marker Systems

Map-based cloning uses the fact that, as distances between the gene of interest and the analyzed markers decrease so does the frequency of recombination. To position any locus in the *Arabidopsis* genome using a large mapping population (1500-2000 individuals) in a 10 kb interval one would need approximately 12,000 well positioned markers. This is feasible with presently available technologies (Peters et al., 2003). However, to perform efficient map-based gene cloning it is important to follow a good strategy by using the appropriate marker systems in the different successive steps of the procedure.

Regional Targeting with High-Volume Marker Technologies

The probability of identifying one or more markers within a specified physical distance of a target gene decreases the closer one gets to the target. Several strategies have been developed that allow one to screen a large number of random molecular markers in a relatively short time and to select just those few markers that reside in the vicinity of the target gene. These methods rely on two principles: (i) the development of high-volume marker technology, which allows

hundreds or even thousands of potentially polymorphic DNA segments to be generated and visualized rapidly from single preparations of DNA and (ii) the use of segregating populations to identify, among these thousands of DNA fragments, those few that are lying in a region adjacent to the gene of interest.

High-volume marker systems such as amplified fragment length polymorphisms (AFLPs) (Vos et al., 1995; Zabeau, 1993) and high throughput single nucleotide polymorphism (SNP) detection (Cho et al., 1999; Gresham et al., 2006) have been used with success in several plant species. These methods allow a global comparison of genomes for genetic differences.

The AFLP method relies on PCR used in conjunction with primers that amplify random sites throughout the genome. Amplification products are analyzed on acrylamide gels to allow visualization of variation among individuals. The fact that multiple AFLP markers can be obtained per PC was exploited by Peters et al. (2004) to develop an AFLP-based genome-wide mapping strategy.

Borevitz and colleagues developed a high-throughput genotyping platform for the identification of SNPs by hybridizing genomic DNA from *Arabidopsis thaliana* accessions to an RNA expression GeneChip (AtGenome1). Using this approach several high heritability development and circadian clock traits were mapped via direct hybridization of mutant DNA to ATH1 expression arrays (Hazen et al., 2005).

By using these high-volume marker technologies, thousands of loci scattered throughout the genome can be assayed in a matter of weeks or months. The next issue, however, is to determine which of the amplified loci lie near the target gene.

In the past few years two approaches have proven to be effective for identifying from a large number of markers the few that reside near the locus

of interest. Both involve the use of genetic stocks that are almost identical, except in the regions flanking the targeted gene.

A first and generally applicable strategy has been referred to as the bulked segregant analysis (BSA) method and relies on the use of segregating populations (Giovanonni et al., 1991; Michelmore et al., 1991). BSA requires the generation of populations of bulked segregants (bulks). When P_1 and P_2 are crossed, the F_2 generation derived from this cross will segregate for alleles from both parents at all loci throughout the genome. If the F_2 population is divided into two pools of individuals with contrasting phenotypes on the basis of screening at a single target locus, these two pools (bulk 1 and bulk 2) will differ in their allelic content only at loci contained in the chromosomal region close to the target gene. Bulk 1 individuals, selected for the recessive phenotype, will contain only P_2 alleles near the target, while bulk 2 plants, selected for the dominant phenotype, will contain alleles from both P_1 and P_2 . Bulk 1 and 2 will both contain alleles from both P_1 and P_2 at loci unlinked to the gene of interest. The use of high volume marker technologies is particularly helpful during the first steps of the map-based cloning procedure, when no information about the location of the gene of interest is available.

The second strategy uses nearly isogenic lines (NILs) but this method can only be applied if inbred lines are generated that differ at the targeted locus (figure 1.16; Alonso-Blanco and Koornneef, 2000). NILs are created when a donor line (P_1) is crossed to a recipient line (P_2). The resulting F_1 hybrid is then crossed back to the P_2 recipient to produce the back-cross₁ generation (BC_1). From BC_1 , a single individual that contains the dominant allele of the target gene from P_1 is selected. Selection for the target gene is performed using markers. The BC_1 individual is subsequently back-crossed again to P_2 , and the cycle of back-cross selection is repeated for a number of generations. In the BC_7 generation most of

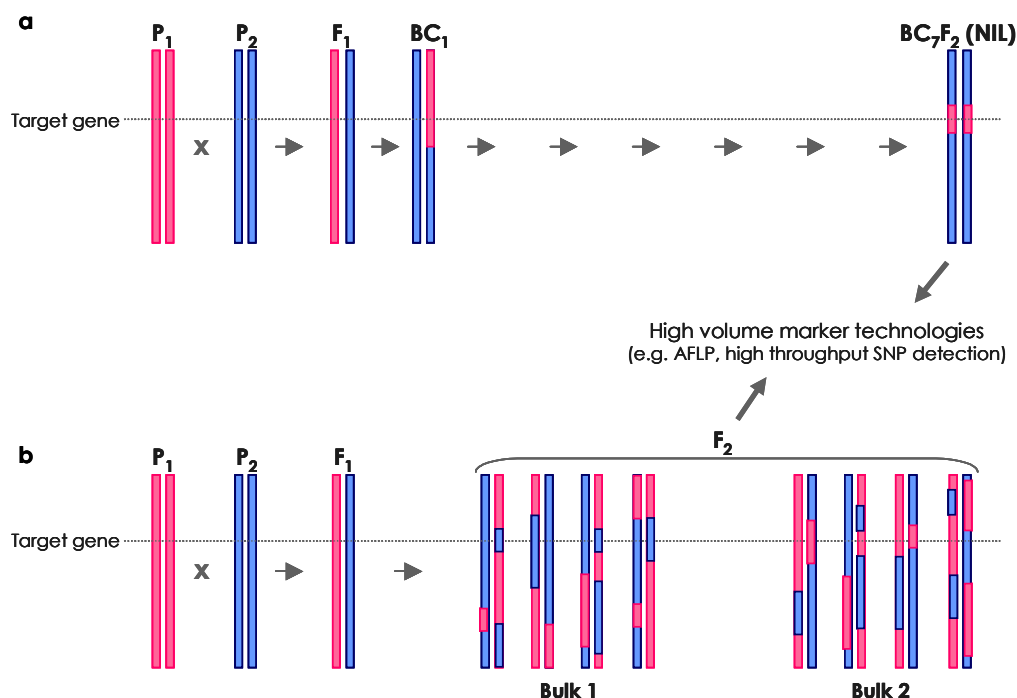


FIGURE 1.16 NEARLY ISOGENIC LINES (NILs) (A) AND BULKED SEGREGANT ANALYSIS (BSA) (B). High volume marker technologies can be used to compare the genetic profiles of pairs of NILs or pairs of bulks to search for DNA markers near the target gene.

the genome will be derived from P_2 , except for a small heterozygous chromosomal segment containing the selected dominant allele, which is derived from P_1 . Lines homozygous for the target gene can be selected from the BC_7F_2 generation. The homozygous BC_7F_2 line is said to be nearly isogenic with the recipient parent P_2 . The genetic profiles of pairs of NILs can be compared using high-volume marker technologies. If the same polymorphisms are observed between the NIL and P_2 and between P_1 and P_2 , they should be derived from a locus tightly linked to the gene of interest.

Fine-Mapping with Individual, Simple and Robust PCR Markers

Once a chromosomal segment has been defined in which the

gene of interest is localized, it is practical to start using individual markers to further narrow down the interval. The *Arabidopsis* InDel polymorphisms (<http://www.arabidopsis.org/Cereon/index.html>) are a good source for such PCR markers but any other easy to use marker (for instance, SSLP) will do. The flanking PCR markers are tested on 1,000-2,000 individuals from a segregating population (e.g. F_2 population). The markers can be analyzed simultaneously, which makes identifying recombinants in the region of interest straightforward. The selected recombinants can subsequently be used to further delimit the area containing the target gene to a region of approximately 5 to 10 kb. Markers necessary for this process can be developed from the available Cereon InDel and SNP collection (Jander et al., 2002).

Outline of a Typical Map-Based Cloning Experiment

Creating a Mapping Population

The first step in the mapping process is to create a mapping population by crossing the mutant to a WT plant of another ecotype (in practice Col and Ler is the most common combination). Usually, it is preferred to clean up the genetic background of the mutant by at least one back-cross cycle. F₁ seeds are sown and, as the plants are growing, it is possible to perform phenotype analysis. Presence or absence of the mutant phenotype in the F₁ generation will suggest whether the mutation of interest is likely to be dominant or recessive. F₂ seeds are collected from self-pollination of the F₁ plants and the phenotype of the F₂ plants is determined, unless the trait can only be scored in the progeny (F₃) seed. A 3:1 segregation of WTs versus mutants indicates a monogenic trait. Mapping resolution can be referred to as the average distance between two recombination breakpoints in a given mapping population. In *Arabidopsis*, a genetic distance of 1% recombination (or 1 centiMorgan (cM)) corresponds, on average, to a physical distance of about 250 kb. However, the ratio between genetic and physical distance is not constant and varies with respect to chromosome position and different mapping populations. The main determinant of mapping resolution is the size of a mapping population. Resolutions in the range of 10 to 40 kb can be obtained in mapping populations of about 1000 plants. Typically, a DNA fragment of this length contains between two and ten genes.

Detecting Linkage

Once the F₂ generation of a mapping cross is available, the mutation is

assigned a rough position on the genetic map by identifying linked genetic markers. Next, the two closest linked markers are defined (two markers that are about 5 cM apart and define a genetic interval containing the mutation).

As already described above high-volume marker technologies, such as AFLP, allow to detect linkage and non-linkage very quickly since a large number of markers dispersed throughout the genome can be visualized simultaneously using just a few PCs. Peters et al. (2004) exploited the fact that multiple AFLP markers can be obtained per PC to develop an AFLP-based genome-wide mapping strategy (figure 1.17). Their strategy establishes linkage to a <6 Mb region by analyzing 20-30 mutant individuals from a segregating population with eight PCs that provide a well-dispersed grid of 85 AFLP markers covering the genome. Further analysis of approximately 120 mutant individuals will typically exhaust the available *Arabidopsis* AFLP map (Peters et al., 2001) and identifies a 200-800 kb region. To reduce the number of PCR reactions (especially when individual PCR markers are used to detect linkage) DNA samples can be pooled for bulked segregant analysis (Michelmore et al., 1991).

Cho et al. (1999) described another genome-wide mapping strategy using SNPs to provide a first rough localization of the mutation on the *Arabidopsis* genome. To obtain a higher resolution (<3.5 cM) only a very limited number of F₂ individuals (28) have to be analyzed. However, it is array-based (Affymetrix, Santa Clara, California, USA) and thus relatively expensive.

Before performing a large-scale screen to select recombinants for further fine-structure mapping, it is essential to confirm that there are no comparable and cloned mutants located in the identified region. This can be checked using the sequence-based map of *Arabidopsis* genes with mutant phenotypes (Meinke et al., 2003). Potential candidates can be checked by

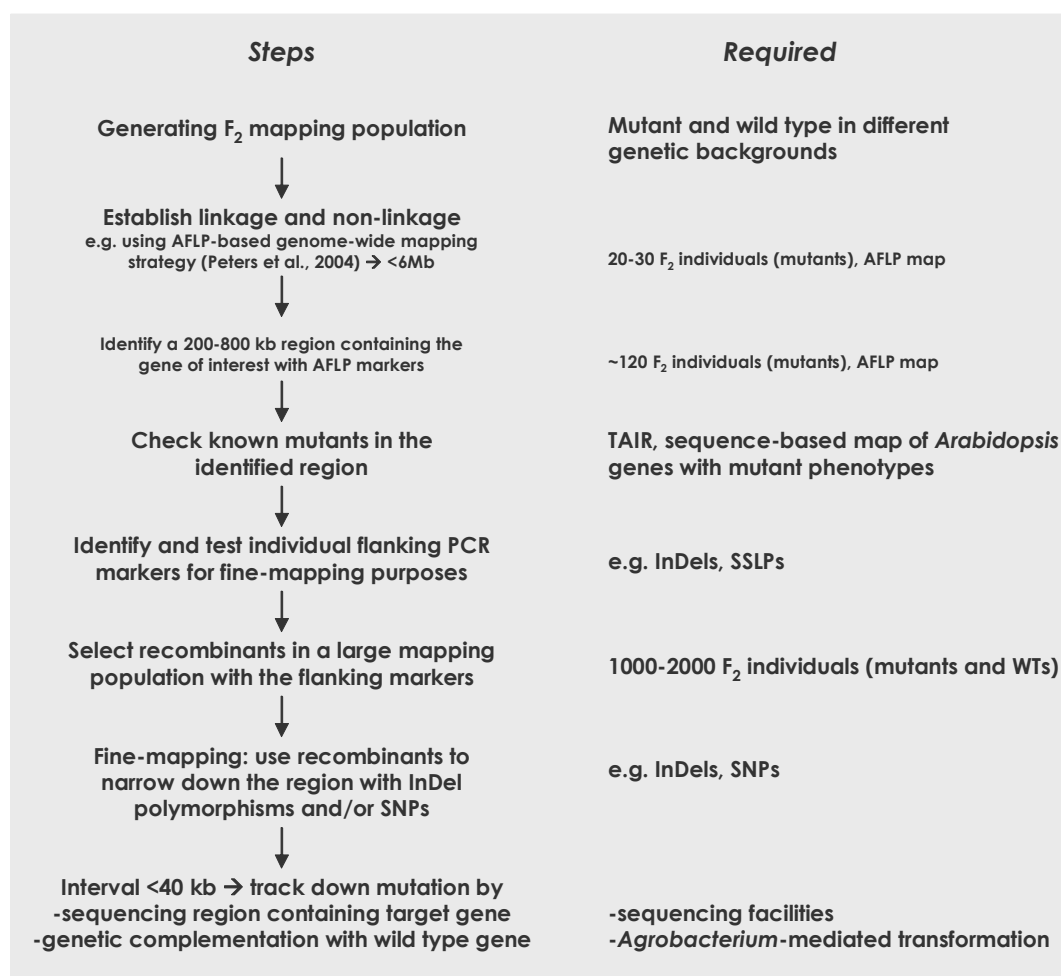


FIGURE 1.17 A MAP-BASED CLONING PROCEDURE (AFTER PETERS ET AL, 2004). The material needed to map a mutation in *Arabidopsis thaliana* is given for each step in the procedure. See Peters et al. (2001) for the AFLP map referred to in steps 2 and 3. For the sequence-based map of *Arabidopsis* genes with mutant phenotypes referred to in step 4, see Meinke et al. (2003). See <http://www.arabidopsis.org/> for the TAIR database; see <http://www.arabidopsis.org/Cereon/index.html> for further details of InDel polymorphisms and SNPs.

performing allelism test crosses. If it is certain that the gene to be cloned is still unidentified fine-mapping can be initiated.

Collecting Recombinants in the Vicinity of the Mutation and Fine-Structure Mapping

Searching a large mapping population for chromosomes with a

recombination in the vicinity of the mutation can be done effectively if two markers are known that are closely linked and that flank the mutation on both sides (figure 1.18). In a large F₂ population (1500-2000 individuals) the majority of plants will have the same genotype for both flanking markers (Lukowitz et al., 2000; Peters et al., 2004). These plants are either not recombinant or, in very rare cases, carry two recombination events. The majority is not informative for mapping the mutation and can be discarded without further analysis. Thus, the bulk of the mapping population is only analyzed with two

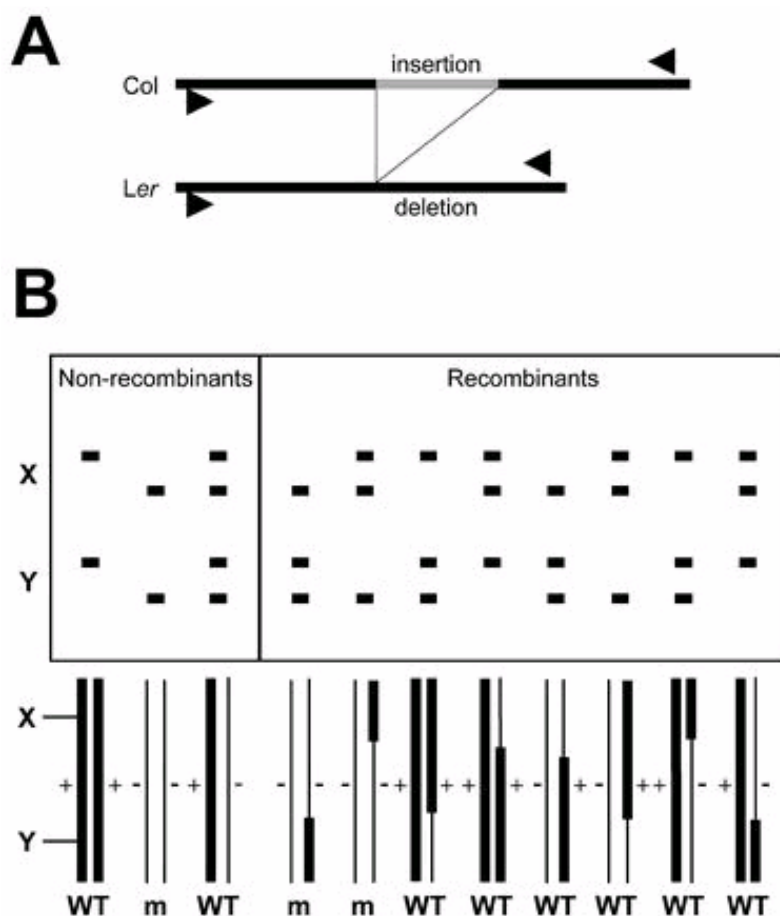


FIGURE 1.18 SCREENING OF A LARGE F₂ POPULATION WITH TWO FLANKING INDEL MARKERS (PETERS ET AL., 2003). **A** Schematic example of an InDel marker. Amplification with the primer pair shown will result in a larger fragment for the Col parent compared to the Ler parent. **B** Two markers (X and Y) represent InDel markers flanking the region of interest from above and below, respectively. Both InDel markers result in a larger fragment for the Col parent compared to the Ler parent. F₂ individuals are subjected to amplification with the primer pair belonging to markers X and Y and the two reactions are run simultaneously on a polyacrylamide gel. F₂ individuals that show a recombination event in the region flanked by X and Y can easily be distinguished from non-recombinants because they produce three PCR fragments instead of two (homozygous F₂ non-recombinants) or four (heterozygous F₂ non-recombinants). Further delineation of the region of interest is an iterative process, using markers localized more inward. The F₂ genotype and phenotype belonging to non-recombinants and recombinants are represented in the bottom panel. The mutant allele is represented by -, whereas the wild-type allele is represented by +.

flanking markers and not considered further. A minority of the F₂ plants will show different genotypes for the two flanking markers. Obviously, these plants carry a chromosome with a recombination close to the mutation and are informative for further mapping.

If the mutation is recessive, viable, and fertile, it is possible to utilize only plants that show the mutant phenotype for fine mapping. Since all these plants should be homozygous mutant for the gene of interest, no further analysis of the F₃ generation is required. Alternatively, all F₂ plants can be included in the mapping population. In this case the genotype with respect to the mutation must be determined for all plants that have been found to carry a recombination event in the vicinity of the gene. Obviously, this often requires analysis of the F₃ generation.

Although a bulked segregant analysis is a very effective way to detect linkage, it usually does not allow determination of the order of closely linked loci on the chromosome. This can only be done by analyzing three-point mapping data. To find flanking markers that can be used to select recombinant chromosomes as described above, it is therefore necessary to examine individual F₂ plants with markers from the region. Therefore, all F₂-individuals first have to be subjected to a screen with the two flanking markers. Since the advantage of high-volume marker technologies is lost at this point, it is advised to use other PCR-based markers. For *Arabidopsis*, as already mentioned before, the Cereon collection of InDels and SNPs can be used for this purpose. InDel polymorphisms flanking the region of interest are identified and primers are

developed to score these flanking markers in a co-dominant way. The identified recombinants can then be used to further narrow down the region of interest.

Identification of the mutated gene

Once an interval of less than 40kb containing the mutation of interest has been determined, this entire region can be sequenced to find the mutation (Jander et al., 2002; Tanksley et al., 1995). Otherwise the number of F₂ plants needed to find recombinants in such a small interval would be very large. Because the sequence of the Col-0 genome is known, one efficient way to sequence the mutant region is to design PCR primers to amplify overlapping segments of about 500 bp spanning the entire 40 kb interval. These segments are then sequenced and assembled, the sequence is compared with that of the WT plant (Col-0 or Ler), and the mutation is identified.

In the case of a mutation in the Ler background, it is necessary to also sequence the Ler WT for comparison at every location where a difference to the WT Col-0 is found.

In the case of a mutation in Col-0, a published sequence is available. However, it is necessary to confirm that any nucleotide that diverges from the published Col-0 sequence was induced by the mutagenesis treatment and is not present in the WT progenitor strain. This is because strain to strain differences exist in Col-0 WT, and even at the high quality standard of the Col-0 sequence, sequencing errors are expected and found (Jander et al., 2002).

The ultimate way to prove that a particular gene is mutated is to complement the mutant phenotype by transformation with the WT gene. This latter approach provides the most direct and definitive evidence. Genetic complementation by *Agrobacterium*-mediated transformation is routine in *Arabidopsis* and has been used in several

instances to confirm that a candidate gene is responsible for a particular phenotypic trait.

1.3.3 REVERSE GENETICS: FROM GENOTYPE TO PHENOTYPE

Rapid progress being made in genome sequencing projects provides raw material for the potential understanding of gene function, so effective reverse genetic strategies are increasingly in demand. Sequence information alone may be sufficient to consider a gene to be of interest, because sequence comparison tools that detect protein sequence similarity to previously studied genes often allow a related function to be inferred. Hypotheses concerning gene function that are generated in this way must be confirmed empirically. Experimental determination of gene function is desirable in other situations as well, for example, when a genetic interval has been associated with a phenotype of interest. In such cases, the functions of genes in an interval can be deduced from the phenotypes of induced mutations. Furthermore, the dissection of gene interactions often requires the availability of a range of allele types. However, most available methods for inferring function rely on techniques that produce a limited range of mutations, are labor-intensive or unreliable, or are limited to species in which special genetic tools have been developed. Just as the discovery of induced mutations led to forward genetics, the introduction of rapid reverse genetic methods can have great impact.

Several general strategies have been used to obtain reduction-of-function or knockout mutations in model organisms, including insertional mutagenesis (Alonso et al., 2003; Bouche and Bouchez, 2001; Long and Coupland, 1998; Parinov and

Sundaresan, 2000; Sussman et al., 2000) and RNA suppression (Chuang and Meyerowitz, 2000; Waterhouse et al., 1998), which have been widely used in plants. Insertional mutagenesis is now largely an *in silico* procedure for *Arabidopsis* researchers, as searchable databases of flanking sequences from T-DNA and transposon insertions are available on-line (Krysan et al., 2002). RNA suppression currently requires considerable manual effort, but it has the potential of reducing expression of repeated genes, which are especially common in plants, including *Arabidopsis*. However, because these techniques rely either on *Agrobacterium* T-DNA vectors for transmission or on endogenous tagging systems, their usefulness as general reverse genetics methods is limited to very few plant species.

Recently, an additional, extremely powerful tool has become available: high-throughput screening for induced point mutations. This allows the identification of induced lesions in a particular gene of interest (targeted, induced lesions in genomes, TILLING) (Colbert et al., 2001). TILLING is a large-scale reverse genetic strategy that uses a high-throughput method based on an endonuclease that preferentially cleaves mismatches in heteroduplexes between WT and mutant DNAs. Subsequent analysis of cleavage products on a sequencing gel allows for the rapid identification of induced point mutations. Unlike insertional mutagenic approaches, which often provide knockout mutations, EMS-mutagenesis-based TILLING yields point mutations that provide an allelic mutant series. For instance, it might identify sublethal alleles, which are especially valuable for analyzing the phenotype of essential genes. TILLING has the potential to identify conditional (temperature-sensitive) alleles in a relatively simple way.

Several TILLING reverse genetics projects are now underway in diverse plant species, including *Arabidopsis* (Till et al., 2003), *Lotus* (Perry et al., 2003), maize (Till

et al., 2004) and *Brassica oleracea* (Gilchrist et al., unpublished results), and in some animal species (Hurlstone et al., 2003; Smits et al., 2004; Stemple, 2004; Wienholds et al., 2003).

Reverse genetics approaches provide a powerful tool to study genes of unknown function and therefore complement the study of genes identified by forward genetic approaches.

1.4 PLANT FUNCTIONAL GENOMICS THROUGH TRANSCRIPTOME ANALYSIS

Reverse genetics represents one category of experimental tools that can be used to decipher the biological significance of the genome sequences from various organisms that are currently accumulating. However, information on both the physical and functional annotation of the genome can also be gained through transcript profiling (Hughes et al., 2001; Shoemaker et al., 2001). In recent years, transcript profiling has become synonymous with gene expression analysis, largely because of the technical difficulties and greater molecular complexity of proteomics and metabolomics (Smith, 2000). Such correlations are acceptable in many cases (Celis et al., 2000), even though the term 'gene expression' is often used to refer more directly to the compendium of gene products that ultimately cause cellular responses, and these are more often than not proteins. In some instances, protein levels are not reflected by alterations in mRNA level, and their activities are often controlled by post-translational modifications (Gygi et al., 1999). This means one has to take care with both experimental design and interpretation if one wishes to extrapolate transcript levels to those of protein (proteomics) or protein activity (metabolomics). Even so, it is generally accepted and there is much experimental evidence to support the statement that

while post-transcriptional events play a role in modulating gene expression the primary level of control is at transcription. Therefore, the accessibility of transcript profiling in recent years has allowed the establishment of various high-throughput methodologies of gene expression analysis. These methodologies differ in their convenience, expense, number of transcripts assayed and sensitivity (Kuhn, 2001).

1.4.1 OVERVIEW OF GLOBAL TRANSCRIPT PROFILING METHODS

1.4.1.1 SEQUENCE-BASED ANALYSIS OF GENE EXPRESSION

In the early nineties, the automatization of the sequencing protocol allowed scientists for the first time to build up a large data collection of so-called expressed sequence tags (ESTs) in model organisms. ESTs are created by sequencing the 5' and/or 3' ends of randomly isolated gene transcripts that have been converted into cDNA (Adams et al., 1991). Generating sequences from cDNA fragments serves two purposes, the discovery of new genes and the assessment of their expression levels in the representative tissue (Ewing et al., 1999; Mekhedov et al., 2000; White et al., 2000). The basis of the approach is that the level of an mRNA species in a specific tissue is reflected by the frequency of occurrence of its corresponding EST in a cDNA library. In this respect the technology is distinct from ratio-based methodologies, such as microarraying, in being immediately quantitative (Bohnert et al., 2001). EST technologies are attractive because they do not rely on established sequence data from the organism under study, and they also fit well with labs already equipped to carry out high-throughput DNA sequencing (Adams et al., 1991). However, even at a few dollars per

sequence the process can be expensive if one desires to progress beyond cursory screening of abundant mRNAs to in-depth analysis (Ohlrogge and Benning, 2000). In addition to the statistical problems in sampling small numbers from a large population (Audic and Claverie, 1997), there are also problems of bias in cloning and cDNA synthesis, though this last problem and others associated with data normalization are not specific to this technology. Auxiliary techniques are available that reduce the amount of sequencing. These include subtraction hybridization (Sargent, 1987) and related methods, representational difference analysis (RDA) (Hubank and Schatz, 1994) and suppression subtractive hybridization (SSH) (Diatchenko et al., 1996). Numerous variants on these technology themes are also available (Bonaldo et al., 1996; Sagerstrom et al., 1997).

Currently, there are nearly 20 million ESTs in the NCBI public collection, more than 4 million of which derive from plants (<http://www.ncbi.nlm.nih.gov/dbEST/>). With the rapid expansion of available EST data (e.g. <http://www.arabidopsis.org/>; <http://www.gramene.org/>; <http://www.medicago.org/>; <http://www.sgn.cornell.edu/>; <http://www.tigr.org/>), opportunities for digital analysis of gene expression will continue to expand.

By developing Serial Analysis of Gene Expression (SAGE), Velculescu et al. (1995) addressed the problem of reducing the costs of traditional EST sequencing, and further modifications have allowed this procedure to handle small amounts of tissue (Datson et al., 1999; Peters et al., 1999). Though a similar sequence-based method to EST analysis, SAGE achieves a cost saving by the concatenation of multiple sequence tags of 10–14 base pairs, prior to cloning. By the size selection of inserts containing 25–50 tags, a comparable reduction of cost or increase in depth of analysis can be achieved over the sequencing of single ESTs. However, this increased efficiency comes at the

price of more extensive sequence reads. Consequently, this technology is best applied to organisms whose genomic sequences are known or that have a substantial cDNA sequence database. Because SAGE tags tend to originate from the 3' portion of transcripts, they are less effectively screened against most cDNA libraries, which are generally sequenced only from their 5' ends. Even with a reference database, because the tags are so short, there can be a redundancy of matches. SAGE is not very convenient for the comparison of many different samples and for the study of less abundant transcripts. SAGE has been used sparingly in plant research, in contrast to the numerous human and yeast studies (<http://www.sagenet.org/>).

As the name suggests, Massively Parallel Signature Sequencing (MPSS) has effectively tackled the problems of EST sequencing with regard to speed and depth of analysis, both through refined molecular biology and, equally important, good quality automation and data management (Brenner et al., 2000a, b). The method, developed at Lynx Therapeutics, Inc., California, is based on the *in vitro* cloning on microbeads of cDNA fragments from an mRNA population using the Megacclone technology (figure 1.19). To achieve this, a collection of oligonucleotide tags is synthesized from a defined set of 4-mers. This collection is so large that when attached to 3' cDNA restriction fragments, tag-cDNA conjugates are obtained with virtually every polynucleotide having a unique tag. These conjugates are then amplified and hybridized to anti-tag sequences specific to separate microbeads. As a result, each microbead hybridizes about 100,000 copies of a single species from the interrogating cDNA. The cDNA fragments can then be sequenced in a flow cell by a method involving type II restriction endonucleases and the tandem ligation of differential adapters, one set of which are fluorescently labeled. A camera with image processing software tracks the

fluorescence of individual beads through the multiple hybridization, ligation, and cleavage steps.

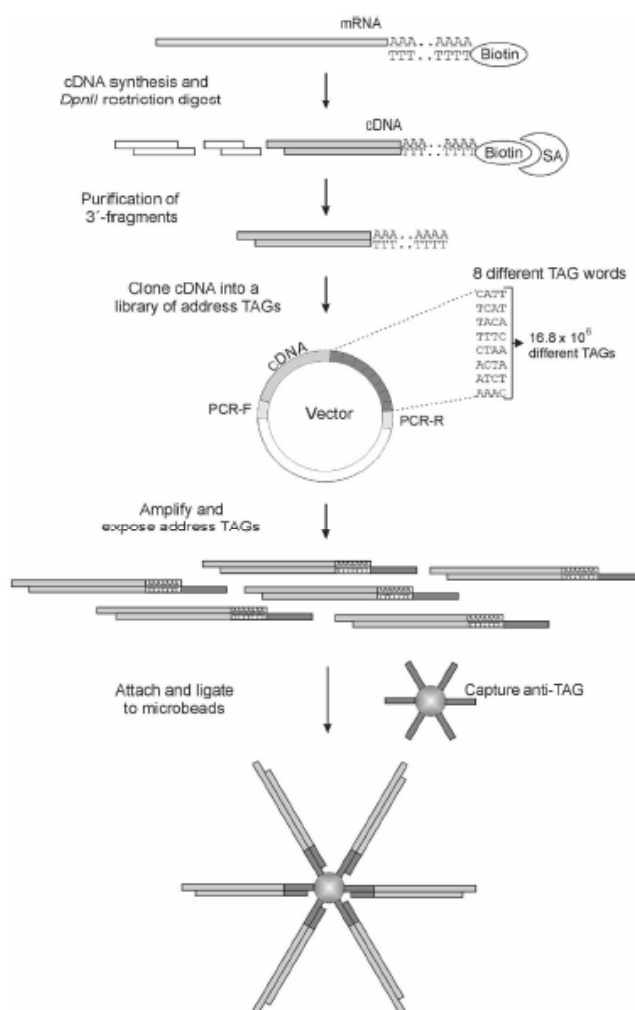


FIGURE 1.19 OVERVIEW OF THE MEGACCLONE TECHNOLOGY

(BRENNER ET AL., 2000A, B). Poly (A) mRNA molecules are converted into double-stranded cDNA molecules, which are digested with *DpnII* and cloned into a specially designed plasmid vector containing a 32 base pair oligonucleotide tag. There are 16.8×10^6 different 32-base sequences available in the reference tag library, and each cDNA clone contains a different sequence. A library of cDNA inserts, along with their adjacent 32 bp oligonucleotide tags are PCR amplified, and the resulting linear molecules are partially treated with an exonuclease to make the 32-base tag single stranded. The 32-base tags at the end of each of the cDNA molecules are hybridized to complementary 32-base tags that are covalently linked to 5 μ m microbeads. There are 16.8 million different complementary tags, each of which corresponds to one of the 16.8 million different 32-base tags. Therefore, for every tag on a cDNA molecule, there will be one bead with a complementary 32-base tag available for hybridization. Each bead contains a vast excess of one particular 32-base complementary tag sequence. Once the cDNA molecules are hybridized to the beads, the nicks are sealed enzymatically. The end product is a micro-bead with approximately 100,000 identical cDNA molecules covalently attached to the surface.

1.4.1.2 FRAGMENT-BASED ANALYSIS OF GENE EXPRESSION

Differential display (Liang and Pardee, 1992; Welsh et al., 1992) uses low-stringency PCR, a combinatorial primer set, and gel electrophoresis to amplify and visualize larger populations of cDNAs representing mRNA populations of interest. Since 1992, there has been a veritable cornucopia of published variations on this theme of differential display (Matz and Lukyanov, 1998). At the time, these methods provided a far more cost effective approach to identifying differentially expressed genes compared to existing subtractive methods. Consequently, they have been used extensively in plant research (Visioli et al., 1997; Sablowski and Meyerowitz, 1998). However, a survey of such studies shows that while differential display has been an effective means of gene discovery it has had less value as a means of gene expression profiling (Sambrook and Russell, 2001). This stems from the inherent nature of the PCR processes used in this and associated methods (McClelland et al., 1995). Because arbitrary primers are used at low annealing temperatures to allow for priming at multiple sites, the amplified products are not solely dependent on the initial concentration of a particular cDNA, but are also a function of the quality of match of the primers to the template. Predictably, under such conditions profiles are highly influenced by the PCR conditions as well as the sample quality (Matz and Lukyanov, 1998). This problem manifests itself in a high percentage of false-positive bands (Sun et al., 1994; Sompayrac et al., 1995), lack of sensitivity (Bertioli et al., 1995), and difficulties with reproducibility (Haag and Raman, 1994; Zhang et al., 1998). The use of longer primers has reduced but not fully alleviated problems of false-positive bands (Zhao et al., 1995; Martin and Pardee, 1999). Similar modifications have been employed to address sensitivity (Ikonomov and Jacob, 1996) and reproducibility

problems (Linskens et al., 1995). However, these changes represent modifications that do not change the inherent problem of arbitrarily primed PCR reactions. In conclusion, differential display methods are relatively cheap and simple means of screening for differentially expressed genes, and are particularly good where the availability of RNA is limited (Renner et al., 1998; Bosch et al., 2000). However, they are not very accurate in quantitatively profiling global levels of gene expression, as illustrated by the number of false-positives generated.

More recently the principles of AFLP (Vos et al., 1995) have been applied to cDNA templates (cDNA-AFLP) (Bachem et al., 1996; Bachem et al., 1998; Goossens et al., 2003; Vandenabeele et al., 2003) and this approach has been used to identify differentially expressed genes involved in a variety of plant processes (Bachem et al., 2001; Breyne et al., 2002; Dellaqi et al., 2000; Durrant et al., 2000; Qin et al., 2000). This technique offers several advantages over more traditional approaches. Of particular importance is the fact that poorly characterized genomes can be investigated in a high throughput manner. Because the stringency of cDNA-AFLP PCR reactions is quite high (which is not the case with differential display) the fidelity of the cDNA-AFLP system allows much greater confidence in acquired data and differences in the intensities of amplified products can be informative (Bachem et al., 1996). In addition, this technique allows a wide variety of tissue types, developmental stages, or time points to be compared concurrently. As with the other profiling methods described here, the sensitivity of cDNA-AFLP is only limited by the ability of cDNA libraries to capture low-abundance transcripts.

Sequencing of cDNA libraries is a more direct and comprehensive approach to gene expression profiling (Adams et al., 1991; Okubo et al., 1992), but this method requires substantial resources for cloning and sequencing, and is less sensitive to

low-abundance transcripts as mentioned above.

Two other fragment-based methods for whole-genome expression analysis were introduced by which cDNAs can be displayed combined with a high sample throughput. A first technology is called GeneCalling, which assays transcript abundance by processing poly-A⁺ RNA-derived cDNA through a restriction digestion utilizing optimized pairs of 6 base pair recognition restriction enzymes. Resultant gene fragments are ligated and end-labeled with 5'-fluorescamine and 3' biotin adapters and subjected to PCR amplification. Subsequently, the fragments are resolved by gel electrophoresis (Bruce et al., 2000; Shimkets et al., 1999). Differentially expressed fragments are identified and confirmed against a virtual digest of the appropriate databases. By combining the sequence knowledge derived from the two restriction enzymes with the length of a given fragment, one is able to 'call' putative gene assignments for each fragment. This information is sufficient for database cross-referencing to determine the identity or novelty (in the case no 'Gene Call' can be made in the database) of a particular gene. This technique proves to be very sensitive, but on the other hand is very laborious and time-consuming.

Another method is named Total Gene expression Analysis or TOGA (Sutcliffe et al., 2000). This approach utilizes sequences near the 3' ends of mRNA molecules to give each mRNA species present in the organism a single identity. This is achieved by exact matching by an equimolar mixture of 48 anchor primers to initiate reverse transcription. One member of this mixture initiates synthesis at a fixed position of each mRNA species, hereby defining a 3' endpoint. By restriction digestion with two enzymes, uniform fragment lengths for each cDNA species are obtained. After PCR amplification, the fragments are visualized and the status of each mRNA can be compared among

samples in an automated fashion. Its identity can be electronically matched with sequences of known mRNAs in databases. This method is sensitive and its single-product-per-mRNA feature allows a near complete coverage of all mRNAs present in a given organism. However, as like 'GeneCalling', this method requires a high input of labor and time.

1.4.1.3 HYBRIDIZATION-BASED ANALYSIS OF GENE EXPRESSION

The principles underlying the hybridization of complementary nucleotide sequences are embodied in the structure of duplexed nucleic acids, and have been exploited experimentally for decades (Gillespie and Spiegelman, 1965). In this time, nucleic acid hybridization has been used in a variety of guises in the quantification of plant RNA levels. In the 1970s, it was used to study sequence complexity (Goldberg et al., 1978). Also in that decade, Southern developed a method using a solid support in hybridization studies of DNA fragments separated by gel electrophoresis (Southern, 1975). This led to a major advance in the analysis of gene expression with the development of northern-transfer hybridization (Alwine et al., 1977). This technique has been immensely useful over the years, though ironically the approach is less global and the method less genomic in nature than preceding solution-based systems. With the availability of nucleotide sequences and clones as physical reagents, hybridization-based approaches now allow for the simultaneous analysis of tens of thousands of genes. Quite literally, one can globally survey the transcription of a plant by hybridization. The interest in this form of transcript profiling has been stimulated by the development of two parallel microarray-based technologies, one based on spotting cDNA fragments (cDNA microarrays) (Schena et al., 1995), the other, the arrayed synthesis of

oligonucleotides (GeneChips) (Lockhart et al., 1996). These two methods have been extensively reviewed in recent years (Duggan et al., 1999; Lipshutz et al., 1999) and both have notable and distinct advantages.

cDNA arrays can be prepared directly from existing cDNA libraries, a large number of which are in the public domain. Thus, fabrication of cDNA arrays is only dependent upon availability of ordered clone collections and appropriate arraying and scanning instrumentation (Clark et al., 1999; Drmanac and Drmanac, 1999; Eisen and Brown, 1999). Once a set of corresponding PCR products has been generated, arrays can be created in multiple versions containing the entire set of available sequences or subsets of sequences resulting in more focussed arrays suitable for specific research applications (regulatory-, pathway-, stage- or response-specific arrays) (Jiao et al., 2003). Such arrays are also useful for reducing a statistical problem of scale (large numbers of features and low number of replications common in microarray experiments). Here, a large array might be used to identify differentially expressed genes of interest, which could then be re-arrayed as a smaller array and used in subsequent experiments. One benefit of this approach is that it can free up resources that can be used to increase experimental replication and thereby increase precision. Another advantage of cDNA arrays is that they can be used in 'two-color' co-hybridization experiments that allow direct comparisons of transcript abundance in two mRNA populations of interest. Although this strategy generates comparative expression ratios instead of measuring absolute expression levels, it is effective for comparative expression profiling and reduces experimental variation that arises in microarray data collected from different chips (Aharoni and Vorst, 2001).

Like cDNA arrays, oligonucleotides can be printed using robotic instrumentation and (once appropriate oligonucleotides have been synthesized) sub-arrays for specific research applications can be fabricated easily. The main limitations in development of oligonucleotide arrays are the costs associated with sequence selection and oligonucleotide synthesis. As these costs continue to decline oligo-based arrays are likely to become more predominant in the near future because they offer a number of important advantages over cDNA arrays. One such advantage is the fact that oligo-based arrays can be fabricated using microfluidic technology, which utilizes light to direct the synthesis of short oligonucleotides onto a suitable matrix (photolithography) (Fodor et al., 1991; Fodor et al., 1993; Pease et al., 1994). Photolithography is particularly useful because it allows for the fabrication of extremely high-density arrays (>300 000 elements/ 1.28 cm²) (Lipshutz et al., 1999). Another important advantage is that the probes in an oligonucleotide array are designed to represent unique gene sequences such that cross-hybridization between related gene sequences (genes belonging to a gene family or genes with common functional domains) is minimized to a degree dependent upon the completeness of available sequence information. Cross-hybridization between homologous sequences continues to be problematic when using cDNA arrays. Furthermore, the array elements in an oligonucleotide array are typically designed to have uniform length, uniform melting temperatures, and to be of uniform concentrations, which can significantly reduce experimental variation and thereby increase statistical power and precision. The primary disadvantage of oligo-based arrays is that oligonucleotide sets can be very expensive because of the extensive sequence data and computational input required for

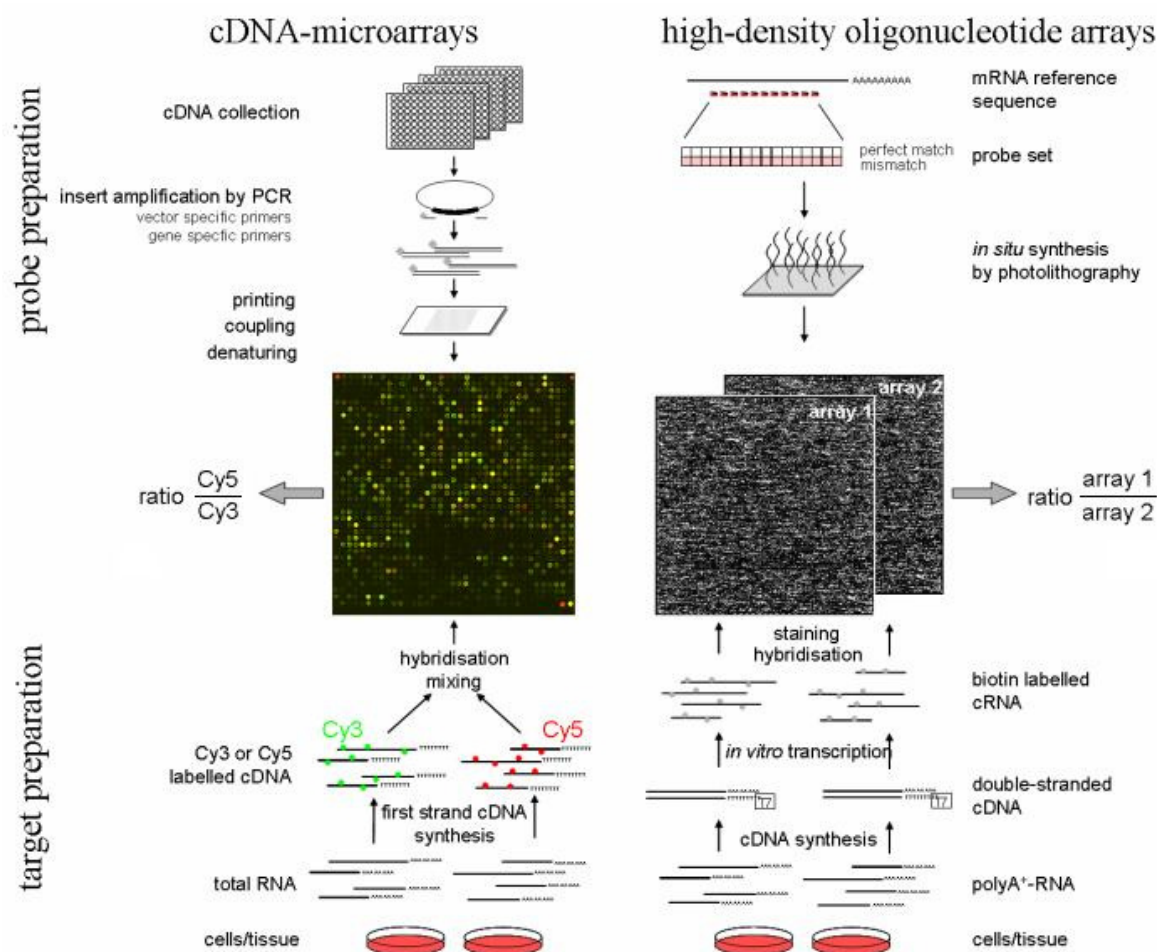


FIGURE 1.20 SCHEMATIC OVERVIEW OF SPOTTED CDNA MICROARRAYS AND HIGH-DENSITY OLIGONUCLEOTIDE ARRAYS (SCHENA ET AL. (1996))

cDNA microarrays:

Array preparation: inserts from cDNA collections or libraries are amplified and the PCR products printed at specified sites on glass slides using high-precision arraying robots. These probes are attached by chemical linkers.

Target preparation: RNA from 2 different tissues or cell populations is used to synthesize cDNA in the presence of nucleotides labeled with 2 different fluorescent dyes (eg: Cy3 and Cy5). Both samples are mixed in a small volume of hybridization buffer and hybridized to the array, resulting in competitive binding of differentially labeled cDNAs to the corresponding array elements. High resolution confocal fluorescence scanning of the array with two different wavelengths corresponding to the dyes used provides relative signal intensities and ratios of mRNA abundance for the genes represented on the array.

High-density oligonucleotide microarrays:

Array preparation: sequences of 16-20 short oligonucleotides (typically 25-mer) are chosen from the mRNA reference sequence of each gene, often representing the unique part of the transcript. Light-directed, in situ oligonucleotide synthesis is used to generate high-density probe arrays containing over 300,000 individual elements.

Target preparation: Total RNA from different tissues or cell populations is used to generate cDNA carrying a transcriptional start site for T7 DNA polymerase. During IVT, biotin-labeled nucleotides are incorporated into the synthesized cRNA molecules, which is then fragmented. Each polymer sample is hybridized to a separate probe array and target binding is detected by staining with a fluorescent dye coupled to streptavidin. Signal intensities of probe array element sets on different arrays are used to calculate relative mRNA abundance for the genes represented on the array.

designing gene-specific oligonucleotide probes. A common limitation of all array approaches is the requirement of significant amounts of RNA for the preparation of fluorescently labelled targets.

Several microarray platforms have been produced based on the *Arabidopsis*

sequence, including first generation arrays produced by a public consortium (The *Arabidopsis* Functional Genomics Consortium or AFGC; Wisman and Ohlrogge, 2000) and a commercial array that included more than 8,000 *Arabidopsis* genes produced by Affymetrix Inc. (Santa Clara, CA) (Zhu and Wang, 2000). The

Affymetrix GeneChip arrays are comprised of sets of 25-base oligonucleotides synthesized *in situ* via a photolithographic process (Lockhart et al., 1996). The most recent generation of the *Arabidopsis* Affymetrix GeneChip can monitor up to 24,000 gene sequences. Affymetrix also offers arrays for a variety of other model organisms including rat, mouse, *Escherichia coli*, *Pseudomonas aeruginosa*, yeast, *Caenorhabditis elegans* and *Drosophila melanogaster*. A second commercial array, produced by Agilent Technologies Inc. (Palo Alto, CA), represents 28,500 *Arabidopsis* genes and is based on the process of ink-jet printing of 60-base probes (Hughes et al., 2001).

Recently, Allemeersch et al. (2005) have assessed the performance of the CATMA microarray designed for *Arabidopsis thaliana* transcriptome analysis and compared it with the Agilent and Affymetrix commercial platforms. The CATMA array consists of gene-specific sequence tags of 150 to 500 bp, the Agilent array of 60-mer oligonucleotides, and the Affymetrix GeneChip of 25-mer oligonucleotide sets. They have matched each probe repertoire with the *Arabidopsis* genome annotation (The Institute for Genomic Research release 5.0) and determined the correspondence between them. Array performance was analyzed by hybridization with labeled targets derived from eight RNA samples made of shoot total RNA spiked with a calibrated series of 14 control transcripts. CATMA arrays showed the largest dynamic range extending over three to four logs. Agilent and Affymetrix arrays displayed a narrower range, presumably because signal saturation occurred for transcripts at concentrations beyond 1,000 copies per cell. Sensitivity was comparable for all three platforms. For Affymetrix GeneChip data, the RMA software package outperformed Microarray Suite 5.0 for all investigated criteria, confirming that the information provided by the mismatch oligonucleotides has no added value. In addition, taking advantage of replicates, they conducted a robust statistical

analysis of the platform propensity to yield false positive and false negative differentially expressed genes, and all gave satisfactory results. The results establish the CATMA array as a mature alternative to the Affymetrix and Agilent platforms (Crowe et al., 2003; Hilson et al., 2004).

1.4.2 MOLECULAR PHENOTYPING THROUGH COMPREHENSIVE TRANSCRIPTOME ANALYSIS

Whole-genome expression profiling is facilitated by the development of DNA microarrays and represents a major advance in genome-wide functional analysis. In a single assay, the transcriptional response of each gene to a change in cellular state can be measured. This change can be a process such as cell division or a response to a chemical or a genetic perturbation (such as a mutation). Because the relative abundance of transcripts is often linked to specific cellular needs, most expression profiling studies to date have focused on the identification of genes that are involved in specific processes of interest, such as systemic acquired resistance, key pathways involved in the circadian clock in *Arabidopsis thaliana*, the analysis of developmental and sex-regulated gene expression in *Caenorhabditis elegans* or the global transcriptional response of *Saccharomyces cerevisiae* to ionic radiation (De Sanctis et al., 2001; Harmer et al., 2000; Jiang et al., 2001; Schenk et al., 2002).

However, the idea that the collection of genome-wide expression profiles in a cell, which consists of thousands of individual observations, can serve as a transcriptional fingerprint and constitutes in this manner a complete and detailed molecular phenotype has only recently begun to receive attention. Provided that the cellular transcriptional response to disruption of different steps in the same pathway is similar, and that there are sufficiently unique transcriptional

responses to the perturbation of most cellular pathways, the systematic characterization of novel mutants could be carried out with a single genome-wide expression measurement.

Using a comprehensive database of reference profiles, the pathway(s) perturbed by an uncharacterized mutation would be identified by simply asking which expression patterns in the database its molecular phenotype most strongly resembles, in a manner analogous to fingerprinting. An important aspect of this approach is that it should be equally effective at determining the consequences of pharmaceutical treatments and disease states. A sufficiently large and diverse set of profiles obtained from different mutants, treatments, and conditions would also result in a relatively comprehensive identification of coregulated transcript groups, allowing additional hypotheses to be drawn regarding the functions of genes based on the regulatory characteristics of their own transcripts (Eisen et al., 1998).

This strategy was elegantly illustrated by Hughes et al. (2000) who analyzed global gene expression in yeast (*Saccharomyces cerevisiae*) by monitoring mRNA expression changes of approximately 6,300 genes under 300 experimental conditions. These data were used to build a reference database or a compendium of expression profiles. The 300 conditions were highly diverse, involving deletion of 279 individual genes from all major functional classes as well as treatment with 13 drugs. Among the deletion mutants, 69 functionally uncharacterized open reading frames (ORFs) were analyzed. These widely divergent experiments led to an alteration in expression of 4,553 genes, or approximately three-fourths of the entire genome in at least one experiment. Moreover, in 96% of the experimental conditions, the expression of at least one gene was significantly altered by more than two-fold and in 41% of the

experiments the expression levels of more than 20 genes were significantly changed. Hughes et al. used their compendium database to identify matching patterns of expression among genes of known function and uncharacterized ORFs. By identifying similar 'fingerprint' patterns, the authors were able to hypothesize, and subsequently verify, the diverse functions of eight orphan ORFs that encode proteins involved in sterol metabolism, cell wall function, mitochondrial respiration or protein synthesis.

To extend this paradigm, they demonstrated that it is possible to identify the target or pathway affected by a therapeutic drug whose mechanism of action is unknown. The expression profile observed in yeast treated with the topical anaesthetic dyclonine closely matched that observed in yeast with a perturbed ergosterol pathway, particularly *erg2Δ* yeast, deficient in Erg2p, a sterol C-8 isomerase. Validation experiments implicated Erg2p as the target of dyclonine. Interestingly, a sequence homology search revealed the likely human homologue to be a σ -receptor, which is not a sterol isomerase, but rather a neurosteroid-interacting protein that regulates potassium channels and presumably accounts for the anaesthetic actions of dyclonine in humans.

The power of the microarray technology, which allows the global analysis of patterns of simultaneous gene expression and moreover the ability to link the database of expression patterns obtained using one microarray to other databases, was also exploited extensively for the classification of human tumours (Ahr et al., 2001; Huang et al., 2003; Sorlie et al., 2001; Sorlie et al., 2003; Van 't Veer et al., 2002). The potential of the approach was first demonstrated by Golub et al. (1999) who determined the RNA expression profiles of a series of acute myeloid leukaemias and acute lymphoblastic leukaemias. A class discovery program automatically identified the distinction between these

leukaemia groups without prior knowledge of the diagnosis. Using this data set, the program was then able to classify new leukaemia cases.

Perou et al. (2001) obtained microarray portraits of a set of 65 surgical specimens of human breast tumours from 42 individuals. The authors selected a subset of 496 genes that showed considerable variation in expression levels between different tumours. When variations in level of expression within the genes were used to order the tissue samples, a striking division of the tumours into oestrogen receptor (ER)-positive and ER-negative categories occurred. Interestingly, the ER-positive group was characterised by the expression of many genes specific to breast luminal cells. The ER-negative tumours could be further divided into several subgroups. A subgroup of six tumours had an expression pattern characteristic of breast basal (myoepithelial) cells, whereas another subgroup was partially characterised by high expression levels of genes associated with overexpression of the *ERBB2* oncogene. On the basis of these results the authors suggested that we are far from a complete understanding of the diversity of human breast tumours.

In yet another study, Pérez-Amador et al. (2001) used DNA microarray analysis for the further molecular characterization of the *dst1* mutant of *Arabidopsis*. The *dst* (downstream) mutants were isolated originally as specifically increasing the steady state level and the half-life of DST-containing transcripts. As such, they offer a unique opportunity to study rapid sequence-specific mRNA decay pathways in eukaryotes. These mutants show a three-fold to four-fold increase in mRNA abundance for two transgenes and an endogenous gene, all containing DST elements, when examined by RNA gel blot analysis. However, they show no visible aberrant phenotype. The authors used DNA microarrays to identify genes with altered expression levels in *dst1* compared with the parental plants. In addition to

verifying the increase in the transgene mRNA levels, which were used to isolate these mutants, they were able to identify new genes with altered mRNA abundance in *dst1*. RNA gel blot analysis confirmed the microarray data for all genes tested and also was used to catalog the first molecular differences in gene expression between the *dst1* and *dst2* mutants. These differences revealed previously unknown molecular phenotypes for the *dst* mutants. Cluster analysis of genes altered in *dst1* revealed new coexpression patterns that prompt new hypotheses regarding the nature of the *dst1* mutation and a possible role of the DST-mediated mRNA decay pathway in plants.

For *Arabidopsis* a special initiative, called the CAGE project, has been set up in 2002 to build a gene expression reference database. A Consortium of European *Arabidopsis* functional genomics centers has teamed up with bioinformatics partners that contribute expertise in microarray data processing, analysis and storage/distribution. A total of 2000 *Arabidopsis* samples are being produced and analyzed under largely standardized conditions. These samples are profiled on CATMA microarrays containing gene specific probes for most *Arabidopsis* genes, to build a compendium of expression profiles. The data will be assessed for statistical significance and submitted to the ArrayExpress database at the European Bioinformatics Institute (EBI). EBI will deliver specific CAGE ontology, and data submission pipelines. The compendium data will be annotated and analyzed for content and confirmation of gene function.

Another example of a compendium of gene expression data is the Genevestigator *Arabidopsis* microarray database and analysis toolbox (Zimmermann et al., 2004), which is a database and Web-browser data mining interface for Affymetrix GeneChip data. Users can query the database to retrieve the expression patterns of individual genes

throughout chosen environmental conditions, growth stages, or organs. Reversely, mining tools allow users to identify genes specifically expressed during selected stresses, growth stages, or in particular organs. Using GENEVESTIGATOR, the gene expression profiles of more than 22,000 *Arabidopsis* genes can be obtained, including those of 10,600 currently uncharacterized genes.

The fundamental advantage of a compendium approach over conventional assays is that it substitutes a single genome-wide expression profile in the place of many conventional (often tedious) assays that measure only a single cellular parameter. Because the compendium approach to determining gene function does not rely at all on the regulatory characteristics of the gene of interest, it has a significant advantage over the widely accepted idea of assessing gene function based on co-regulation. Furthermore, the same compendium used to characterize mutants can also be used to characterize other perturbations, including treatments with pharmaceutical compounds, and potentially disease states as well.

CHAPTER TWO

ISOLATION OF TEMPERATURE-SENSITIVE GROWTH MUTANTS OF *ARABIDOPSIS THALIANA*

We designed a screening procedure to isolate temperature-sensitive mutations that affect plant growth and development. By applying our screening on an EMS mutagenized population of *Arabidopsis*, we obtained a collection of twelve growth mutants, each of which potentially corresponds to an unknown (essential) plant gene.

A general phenotypic analysis allowed us to classify the twelve mutants into two distinct categories. A first category consists of eight mutants (*dgr1* to *dgr8*; *defective in growth*), which are inhibited in overall plant growth. A second category comprises four mutants (*dlr1* to *dlr4*; *defective in lateral root formation*) that show defects specifically in the process of lateral root formation.

Since we aimed to eventually identify novel essential genes involved in plant growth and development, *dgr1* and *dgr2* appeared, based on their specific phenotypes, to be the most representative candidates to focus our further research on. Whereas both mutants are inhibited in overall plant growth and development, *dgr1* was more specifically selected for its irreversible phenotype and *dgr2* as a representative for a subclass of mutants with a reversible phenotype.

In first instance, our study demonstrates that temperature-sensitive mutants provide a powerful tool for studying gene function *in vivo*. Secondly, we were able to isolate twelve interesting mutants, for which further phenotypic and molecular genetic characterization is expected to provide a diversity of information, potentially useful for further understanding of the biological mechanisms underlying plant growth and development.

2.1 INTRODUCTION

Since the publication of the *Arabidopsis* genome sequence (The *Arabidopsis* Genome Initiative, 2000), which was the first to be completed among flowering plants, the fundamental question in the field of plant growth and development has been how many and which of these thousands of genes are essential to produce a viable and fertile plant with a normal phenotype. To determine the biological significance of a gene product and its relevance to plant growth and development both forward genetics and saturation mutagenesis on a genomics level are required. When gene functions are not redundant, this objective can be met by investigating which gene disruptions result in a mutant phenotype. Using T-DNA insertion mutagenesis, Tzafrir and co-workers (2004) recently constructed a first large-scale dataset of essential *Arabidopsis* genes, required for normal embryo development. Although the identification of essential genes is straightforward using regular knock-out mutants, the study of gene function *in vivo* is often problematic due to the resulting lethal phenotypes.

Screening for conditional lethal mutants provides a means to genetically study mutations in essential genes. The potential to control the activity of individual genes reversibly and quantitatively has opened up new perspectives for the in-depth analysis of gene function. Temperature-sensitive (ts) mutations are the most widely used type of conditional mutations in which there is a marked drop in activity of the gene product when the gene is expressed above (or below) a certain (restrictive) temperature. Below (or above) this temperature, at so-called permissive temperatures, the phenotype of the mutant is similar to that of the WT (WT). This way, gene function can be studied at the restrictive temperature, while mutant lines can be maintained and propagated at the permissive temperature.

Although there are other excellent inducible systems for gene expression (Gossen and Bujard, 1992; Gossen et al., 1995), ts mutants have several unique advantages over such systems, including fast temporal response, high reversibility and the applicability to any tissue type or developmental stage of an organism upon which temperature changes can be imposed. Such organisms include viruses (Shenk et al., 1975), prokaryotes (Fangman and Novick, 1968; Carl, 1970; Wechsler and Gross, 1971; Kirschbaum et al., 1975), yeasts (Hartwell, 1967; Hartwell et al., 1974), nematodes (Hirsh and Vanderslice, 1976; Miwa et al., 1980; Cassada et al., 1981; Kempthues et al., 1988a, Kempthues et al., 1988b), insects (Suzuki, 1970; Suzuki, Kaufman and Falk, 1976; Hall, 1978a; Hall and Greenspan, 1979), amphibia, fish (Nechiporuk et al., 2003; Parichy et al., 2003), and plants.

In the past, several ts mutants have been isolated in plants and most of them show normal phenotypes at low temperature and mutant phenotypes at high temperature. The *deficiens* mutant of *Antirrhinum majus*, for instance, was used in the functional analysis of homeotic genes (Schwarz-Sommer et al., 1992; Zachgo et al., 1995). Other examples include carrot variants unable to complete embryo development (Breton and Sung 1982; Schnall et al. 1988), auxin-auxotrophic mutants of *Nicotiana plumbaginifolia* (Fracheboud and King, 1991) and the *argentina* mutant of *Zea mays*, being delayed in normal development of photosynthetic cell types (Langdale et al., 1987).

In *Arabidopsis*, the ts mutant *rsw1*, which harbours a point mutation in the catalytic subunit of cellulose synthase, has contributed greatly to the molecular analysis of cellulose biosynthesis (Arioli et al., 1998; Peng et al., 2000, Lane et al., 2001). Other examples of *Arabidopsis* ts mutants include the *add* mutants (Pickett et al., 1996), defective in early shoot development, and the *srd* and *rrd* mutants, disturbed in the redifferentiation of shoots from root explants or roots from

hypocotyl explants, respectively (Ozawa et al., 1998; Sugiyama, 1999; Yasutani et al. 1994). Only in rare cases, has high temperature been permissive and low temperature restrictive. Two examples are sweetclover mutants defective in chlorophyll production (Bevins et al., 1993) and the *Arabidopsis* mutant, *fab2*, which overproduces the fatty acid stearate (Lightner et al., 1994).

In this report we describe a screening procedure to isolate *ts* mutations that affect plant growth and development. By applying our screening on an EMS mutagenized population of *Arabidopsis*, we obtained a collection of twelve growth mutants, each of which potentially corresponds to a novel (essential) plant gene.

Initially, a general phenotypic analysis allowed to classify the twelve mutants into two distinct categories. A first category consisted of eight mutants (*dgr1* to *dgr8*; defective in growth) inhibited in overall plant growth. The observed phenotype was either reversible (for *dgr2* to *dgr8*) or irreversible (for *dgr1*), which, at the higher temperature, was reflected in growth arrest or lethality, respectively. A second category comprised four mutants (*dlr1* to *dlr4*; defective in lateral root formation) that showed defects specifically in the process of lateral root formation. Of these, two mutants (*dlr1* and *dlr2*) were not able to form lateral roots at all, whereas the remaining two mutants (*dlr3* and *dlr4*) produced smaller lateral roots compared to WT plants. At a permissive temperature of 18°C the phenotypes of the *ts* mutants were comparable to those of WT plants.

In a next stage, a more detailed phenotypic characterization was focussed on two specific mutants, *dgr1* and *dgr2*. Both mutants were selected from the first class, since we were particularly interested in isolating genes essential for the growth and development of the whole plant. *Dgr1* was furthermore chosen for its irreversible phenotype, whereas *dgr2* was representative for the subclass of mutants

with a reversible phenotype. A Mendelian segregation analysis was performed on both *dgr1* and *dgr2* as a first step towards the isolation of the corresponding genes by positional cloning.

2.2 EXPERIMENTAL PROCEDURES

2.2.1 PLANT MATERIAL

The plant material used in the present study was derived from the Col strain (Col-0) of *Arabidopsis thaliana* (L.) Heynh that is referred to as the WT in this paper. Chemical mutagenesis was performed with ethyl-methane sulfonate (EMS) (15µM) for 24 hours at 24°C on a population of 200,000 seeds (M_1 generation). The M_2 progeny was screened for *ts* mutations. Subsequently, heritability of the observed phenotype was tested with M_3 plants produced by self-fertilization of M_2 plants of putative mutant lines. The M_3 seeds are used to characterize the mutants identified.

2.2.2 GROWTH CONDITIONS

Seeds were sterilized by a ten-minute incubation with a 1.5% sodiumhypochlorite solution (Merck, Germany). Sterile seeds were sown on MS germination medium (0.44% MS, incl. micro- and macro-elements, no vitamins; 0.05%MES; 1% sucrose; 0.8% agar; pH5.7) (Murashige and Skoog, 1962) in square Petri dishes (120mm x 120mm, Greiner bio-one, Frickenhausen, Germany). After sowing, Petri dishes were placed vertically in a growth chamber under continuous light provided by V.H.O. Sylvania fluorescent white tubes at an irradiance of 50 µmol photons m⁻².s⁻¹ (Osram, Milan, Italy) or by Lumilux Cool White light tubes (Osram, Milan, Italy) at an irradiance of 50 µmol photons m⁻².s⁻¹.

2.2.3 SCREENING FOR *ARABIDOPSIS* TEMPERATURE-SENSITIVE GROWTH MUTANTS

(WORK OF ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN)

EMS mutagenized seeds were subdivided in M_1 batches consisting of 800 individual seeds. Seedlings were grown at the permissive temperature of 18°C and were allowed to self-pollinate, which resulted in the corresponding M_2 seed batches.

For 150 M_2 batches, 1000 seeds of each batch were arbitrarily chosen and grown *in vitro* for seven days at the permissive temperature. This was done by taking aliquots of 25 milligrams of seeds (± 1000 seeds) from each M_2 seed batch and washing the seeds for 30 minutes with water and subsequently for five minutes with 70% ethanol. The sterile seeds were sown on sterile strips of filter paper, which were subsequently placed onto MS germination medium. Each collection of 1000 seeds, originating from a single M_2 batch, was divided over four Petri dishes, which were placed at the permissive temperature for seven days.

Seedlings were subsequently screened for deficiencies in lateral root formation at the restrictive temperature of 28°C in order to isolate mutants with *ts* defects in potentially essential genes involved in growth and developmental processes. This was done by transferring the strips of filter paper supporting the one-week old seedlings to new sterile Petri dishes containing the same MS medium supplemented with 10^{-6} M NAA (1-naphthaleneacetic acid, NAA, Sigma, N-0640, Bornem, Belgium) after autoclaving. Those Petri dishes were placed at the restrictive temperature. After four days, the seedlings were examined for the presence or the absence of lateral roots. The candidate mutants failing to form normal lateral roots at the restrictive temperature were rescued by transferring them back to the permissive temperature on fresh sterile medium. The seedlings that recovered were transferred to soil to allow self-fertilization and seed production. In order

to confirm the *ts* phenotype, the resulting M_3 seeds of these rescued plants were re-screened, using the same procedure.

2.2.4 PLANT PHENOTYPIC ANALYSIS (WORK OF ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN)

On soil, plants were sown and grown on a mixture of leaf mould, clay and sand in eight-centimeter pots and were watered daily with tap water.

In hydroponic culture, seedlings were grown on Hoagland nutrient solution diluted four times with demineralized water (six plants per liter of solution). No aeration of the solution was needed. The solution was replaced after the first month of culture and subsequently every two weeks.

Under both conditions, mutant and WT plants were grown at the permissive temperature in a growth chamber under continuous light at an irradiance of $50 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ provided by V.H.O. Sylvania fluorescent white tubes (Osram, Milan, Italy) until the first flower primordia were visible at the center of the rosette. They were then transferred to the restrictive temperature and examined two weeks later.

2.2.5 IN VITRO GROWTH STUDIES

2.2.5.1 AUXIN STUDIES

Ten mutant seeds together with five WT seeds were sown on one Petri dish. For each mutant line three Petri dishes were sown. After sowing, the Petri dishes were placed vertically at the permissive temperature for seven days. Subsequently, plants on the first Petri dish were transferred to fresh MS germination medium without NAA and placed at the restrictive temperature. Plants on the second Petri dish were transferred to fresh MS germination medium supplemented with 10^{-6} M NAA and were subsequently placed at the restrictive temperature.

Plants from the third Petri dish were transferred to fresh MS germination medium supplemented with 10^{-6} M NAA and stayed at the permissive temperature. After four days, pictures were taken using a CAMEDIA C-3040 zoom digital camera (Olympus, Tokyo, Japan).

2.2.5.2 CALLUS INITIATION EXPERIMENTS (WORK PARTIALLY DONE BY ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN)

Using the following procedure, calli were initiated from root explants (Valvekens et al., 1988). Sterile seeds were sown on MS medium and placed for two weeks at the permissive temperature. Subsequently, the roots of the seedlings were excised and transferred to fresh MS medium supplemented with 0.05 mg/L 6-furfurylaminopurine (kinetin, Fluka Chemika, Germany) and 0.5 mg/L 2, 4-dichlorophenoxyacetic acid (2,4-D; Sigma, St. Louis, MO 63178, USA). Half of the explants were then incubated at the permissive temperature and the other half at the restrictive temperature under the light conditions specified above. After six to eight weeks, the explants were examined for the presence of calli and pictures were taken using a CAMEDIA C-3040 zoom digital camera (Olympus, Tokyo, Japan).

2.2.5.3 GROWTH ANALYSIS OF THE PRIMARY ROOT

On each Petri dish fifteen WT and fifteen mutant seeds were sown. For each mutant line, two Petri dishes were sown and placed vertically at the permissive temperature for seven days. Subsequently, the two plates were placed vertically for four days at the restrictive temperature. Eventually only one plate was transferred back to the permissive temperature for five days. After sixteen days, whole plates were photographed using a CAMEDIA C-3040 zoom digital camera (Olympus, Tokyo, Japan) and scanned on a color copier CLC-iR C3200

(Canon, Tokyo, Japan). During these sixteen days, the length of the main root was marked daily by scratching a horizontal line in the Petri dish at the height of the main root tip with a razor blade. Plate scans were analyzed using the image analysis software program Scion Image Beta 4.03 (Scion Corporation, Maryland, USA), the difference in length between each mark was determined. The growth rate for mutant or WT plants was then calculated by dividing the mean length difference of fifteen mutant roots or fifteen WT roots, respectively, by the time elapsed between the two marks.

2.2.5.4 ANALYSIS OF CORTICAL CELL LENGTHS IN THE PRIMARY ROOT

Mutant test plants were grown on MS medium for seven days at the permissive temperature and then transferred to the restrictive temperature for various periods of time (between two and five days), depending on the mutant line. Simultaneously, mutant control plants and WT plants were grown continuously at the permissive temperature and under the same conditions as the mutant test plants, respectively. Subsequently, the seedlings were harvested and fixed for two hours using 4% paraformaldehyde (Merck, Germany) in 20mM Pipes buffer (Merck, Germany) + 1mM CaCl_2 (Merck, Germany) + 0.5% Triton X-100 (Sigma, St. Louis, MO 63178, USA). After fixation seedlings were washed three times for ten minutes with 20mM Pipes buffer + 1mM CaCl_2 + 0.5% Triton X-100. Plants were subsequently mounted on microscope slides in glycerol.

Roots were viewed under DIC (differential interference contrast) optics with a 20x objective (Leica DMLB, Germany). A focal plane was chosen by visualizing the quiescent center. For each root a series of overlapping images of the two cortical cell files in this focal plane were captured using a charge-coupled device camera (Zeiss Axiocam, Germany), fitted to the microscope. The series of images

continued until cell length exceeded the width of the video image (when average cell size was approximately 100 μm). Composite images were created (PanaVue ImageAssembler, Canada) and used to measure the length of every cell in each cortical cell file that could be followed over most of its length (ImageJ software). The position of each cell was defined by its midpoint relative to the quiescent center.

For both mutant plants grown at the restrictive temperature, mutant plants grown at the permissive temperature and WT plants grown at both temperatures, cell lengths from the cortical cell files of five roots were combined and then smoothed and interpolated using the procedure described by Beemster and Baskin (1998). The smoothing procedure fitted a series of overlapping, independent polynomials to the data. First, a given number of data points was selected symmetrically around the first desired x , x being the position of each cortical cell relative to the tip of the root. The number of data points used for each fit was determined as the number of cells between the inflection point where cell length started to increase and the most basal data point and ranged from 15 to 25. The value of x was then increased by 25 μm and the process was repeated. For positions at the beginning and end of the series, the data were not symmetric around x but contained the same number of points. Therefore, a second step was introduced, in which the equally spaced data obtained from step 1 were smoothed further using a similar procedure. The length of the interval used for step 2 was defined as the distance between the same inflection point and the quiescent center and ranged between 450 and 550 μm . The iteration in step 2 was repeated until the change in cell length between successive iterations became smaller than 0.5 μm for all positions. The smoothing algorithm was implemented as a macro for the software program Excel (version 2002; Microsoft Corporation).

2.2.6 BACK-CROSSING AND MENDELIAN SEGREGATION ANALYSIS

Mutant M_3 -plants (grown on soil) were back-crossed with WT Col plants (grown on soil) at the permissive temperature. This was done under a binocular (Leica MZ 16F, Germany) using tweezers to transfer pollen from the WT Col plant to the stamen of the mutant M_3 -plant. Through selfing of the resulting BC_1 -plants BC_1F_2 -seed batches were obtained for each mutant line. The segregation of the BC_1F_2 -families was tested for a 3:1 ratio of WT to mutant phenotypes. Therefore, for each family at least 50 plants were grown for seven days *in vitro* on MS medium at the permissive temperature and subsequently transferred to the restrictive temperature. For each mutant line a number of plants showing a mutant phenotype at the restrictive temperature were rescued by transferring them to soil and back to the permissive temperature. Self-pollination of these mutant plants led to the production of mutant seed batches for each mutant line of which the genetic background was purified for fifty percent. The homogeneity of the resulting mutant seed batches was tested by growing for each batch at least 50 plants for seven days at the permissive temperature, subsequently transferring these plants to the restrictive temperature, and checking the consistency of the mutant phenotype.

2.3 RESULTS

2.3.1 IDENTIFICATION OF TEMPERATURE-SENSITIVE GROWTH MUTANTS OF *ARABIDOPSIS THALIANA* (RESULTS OBTAINED BY ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN)

We screened an EMS mutagenized population of *Arabidopsis* (Col accession) for deficiencies in lateral root formation, when grown at a restrictive temperature of 28°C in the presence

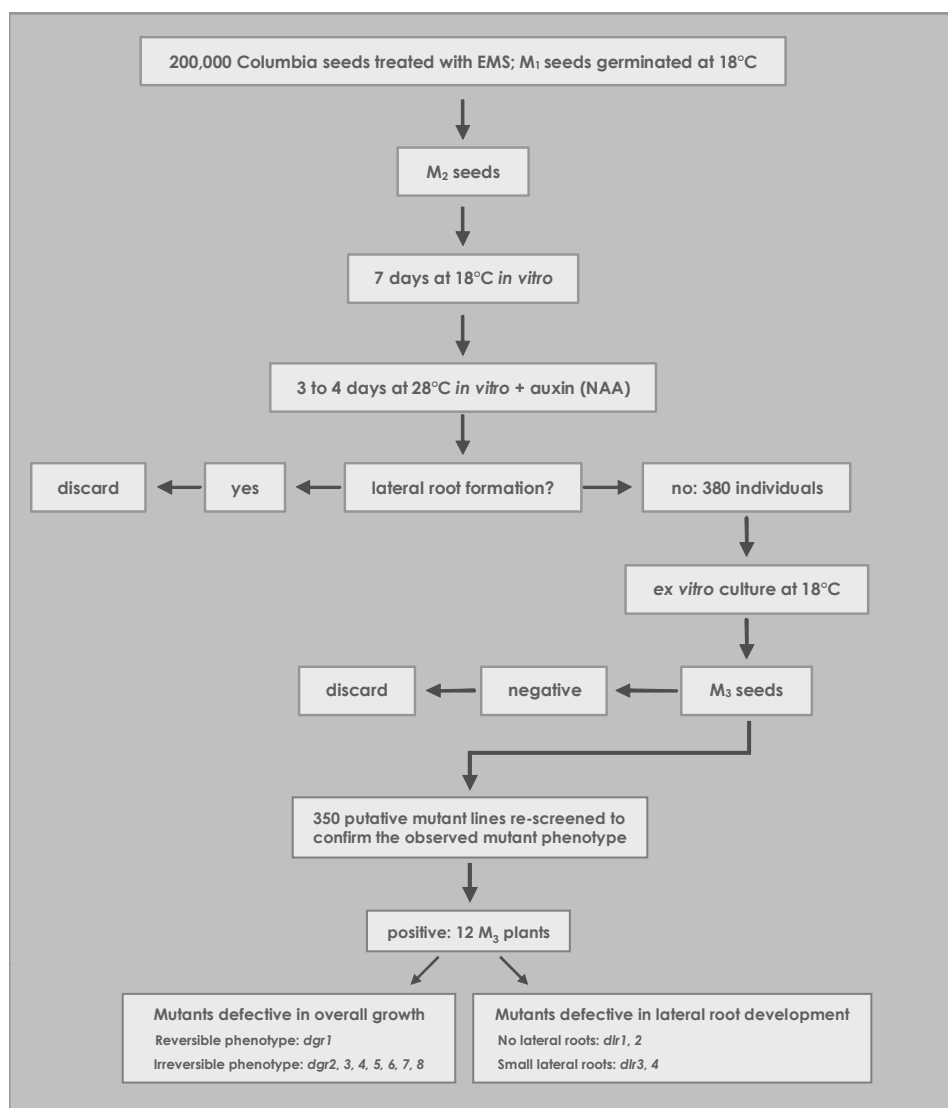


FIGURE 2.1 SCREENING PROCEDURE USED TO ISOLATE TEMPERATURE-SENSITIVE GROWTH MUTANTS OF *ARABIDOPSIS THALIANA*

of auxin (figure 2.1). By selecting for defects in the generation of lateral roots, we aimed to isolate growth mutants in a reasonably fast and straightforward manner. Furthermore, we used exogenous auxin during the entire screening procedure, since it is expected to be easier to distinguish mutant plants that are not able to form any lateral roots from WT plants producing many lateral roots in the presence of auxin.

Initially, 380 candidate mutants were isolated and transferred to the

permissive temperature (18°C) to allow self-pollination. The M₃ progenies were re-screened, using the same screening procedure, in order to confirm the heritability of the ts phenotype. Most of the M₃ plants either showed a WT phenotype or did not germinate well. In the end, we obtained twelve ts lines, disrupted in the production of lateral roots at the restrictive temperature (table 2.1).

Based on the mutant phenotype *in vitro* at the restrictive temperature, mutants were classified into two distinct categories. A first category consisted of

eight mutants (*dgr1* to *dgr8*) exhibiting an overall growth arrest and thus complete inhibition of the development of both shoot and root organs. At the restrictive temperature, *dgr2*, 3, 4, 5, 6, 7 and 8 showed a reversible phenotype if plants were transferred back to the permissive temperature within four days. *Dgr1*, on the other hand, showed an irreversible growth arrest followed by lethality within four days at the restrictive temperature (figure 2.2; B and C).

A second class comprised four mutants (*dlr1* to *dlr4*) with deficiencies in the production of lateral roots. Two mutants, *dlr1* and *dlr2*, were not able to form any lateral roots at the restrictive temperature, although growth of all the other plant organs proceeded normally (figure 2.2; D and E). The other two mutants, *dlr3* and *dlr4*, produced smaller lateral roots compared to the WT (figure 2.2; F).

At the permissive temperature of 18°C the phenotype of most mutants was comparable to the WT phenotype. However, three mutants (*dgr3*, *dgr4*, and *dgr8*) were significantly smaller than WT plants at the permissive temperature.

2.3.2 GENERAL PHENOTYPIC ANALYSIS OF THE TWELVE ISOLATED TEMPERATURE-SENSITIVE MUTANTS

To get a better insight in the different phenotypes of the *dgr* and *dlr* mutants, we initially carried out a general phenotypic characterization. To this end, we first examined the effect of exogenous auxin on the mutant phenotypes. Next, the phenotypes of the mutants were studied during growth on soil and in a hydroponic system. Furthermore, to check whether cell proliferation was affected, we tried to initiate callus from mutant root explants. Finally, an analysis of the growth rate of the main root was performed.

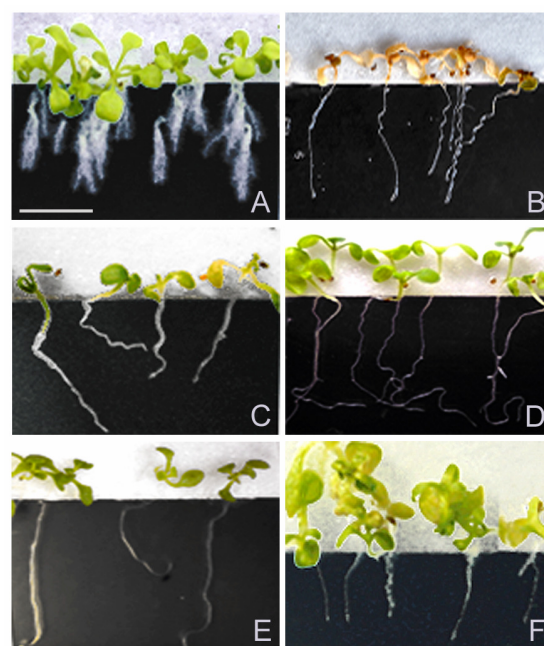


FIGURE 2.2 ARABIDOPSIS *ts* PHENOTYPES OBSERVED AFTER SCREENING FOR DEFICIENCIES IN LATERAL ROOT FORMATION (RESULTS OBTAINED BY ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN). Plants were grown for seven days at the permissive temperature (18°C) before transfer to the restrictive temperature (28°C) (in the presence of 10^{-6} M auxin). Pictures were taken after four days at 28°C. (A) Col WT. Based on the *in vitro* phenotype at 28°C (B) *dgr1* was assigned to a first category of mutants (inhibition of overall growth and development) since it shows complete growth arrest followed by lethality after four days at 28°C. (C) *dgr2* also belongs to this first class but differs from *dgr1* because the growth arrest can be rescued by transferring them back to 18°C within four days at 28°C. (D) and (E) illustrate the phenotypes of *dlr1* and *dlr2*, belonging to a second category (defects in lateral root formation, while growth of shoot and main root is normal). These two mutants are not able to form any lateral roots at 28°C. Mutant *dlr4* (F), also disrupted specifically in the process of lateral root formation, is able to initiate lateral root primordia, but the produced lateral roots are smaller compared to the WT. (bar=1cm)

2.3.2.1 AUXIN DEPENDENCY OF THE TEMPERATURE-SENSITIVE PHENOTYPES

We examined the phenotype of the twelve *ts* mutants both in the presence and in the absence of 10^{-6} M exogenous auxin to investigate whether this phytohormone, initially used in the standard screening procedure, has an effect on the mutant phenotype *in vitro*.

As expected, in the presence and in the absence of exogenous auxin, generation of lateral roots was inhibited in all mutants (figure 2.3; A and B). Our results also show that exogenous auxin

does not induce a changed phenotype in comparison to the WT at the permissive temperature (figure 2.3; C).

2.3.2.2 MUTANT PHENOTYPES AT THE VEGETATIVE AND REPRODUCTIVE DEVELOPMENTAL STAGE

(RESULTS OBTAINED BY ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN)

To further characterize the phenotype of the isolated mutants, they were grown on soil and in a hydroponic system. The plants were initially cultivated at the permissive temperature and subsequently transferred to the restrictive temperature when the first flower buds appeared. This developmental stage was chosen to examine the *ts* defects during the reproductive stage. Two weeks after the transfer to the restrictive temperature, plants were examined and photographed.

Grown on soil, *dgr* mutants showed an inhibition of growth and development at the restrictive temperature but not at the permissive temperature (figure 2.4; B-E). For some mutants, this inhibition was quite severe as in *dgr1* (figure 2.4; B, G, and H), where rapid necrosis was observed, and in *dgr6* (figure 2.4 C, and M), where growth at the higher temperature induced flower abortion. As also observed *in vitro*, *dgr* phenotypes were reversible when the period at the restrictive temperature did not last longer than four days. However, *dgr1* was an exception by showing overall growth arrest followed by lethality. In contrast, *dlr* mutants grown on soil grew better at the restrictive temperature than at the permissive temperature (figure 2.4; A and F).

Cultivation of the mutants in a hydroponic system allowed us to observe the root system and to compare it with WT roots. At the restrictive temperature *dgr* mutants like for example *dgr6*, showed disruptions in the formation of secondary roots but also the growth of the root and shoot was affected (figure 2.4 K, and L; figure 2.5). The *dgr6* mutant turned

necrotic at the apex and did not bolt at 28°C, whereas flowers of *dgr3*, were aborted and sterile (figure 2.4 I, and J).

2.3.2.3 EFFECT OF THE MUTATIONS ON CELL PROLIFERATION

(RESULTS PARTIALLY OBTAINED BY ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN)

To determine whether the *ts* mutations affect fundamental activities required for cell proliferation, the initiation of callus from root explants was examined in each mutant both at the permissive and at the restrictive temperature.

Dgr mutants were unable to initiate calli at the restrictive temperature (figure 2.6). At the permissive temperature, normal callus formation was observed compared to the WT. On the other hand, all *dlr* mutants initiated callus both at the permissive and the restrictive temperature (data not shown).

2.3.2.4 IN VITRO GROWTH ANALYSIS OF THE PRIMARY ROOT MERISTEM

For the *dgr* mutants, we performed an *in vitro* growth analysis of the root apical meristem in order to further characterize their growth behavior at the restrictive temperature. Since the main root growth is not affected in the *dlr* mutants, these were not included in the experiment. For the eight *dgr* mutants a growth decrease of the main root and eventually a complete arrest were observed after transfer to the restrictive temperature. The root growth defects were reversible for all mutants except for *dgr1* (figure 2.7). In this mutant, lethality occurred after four days at the restrictive temperature. In addition, *dgr3*, *dgr4*, and *dgr8* showed already at the permissive temperature significant growth rate differences compared to WT roots (data not shown).

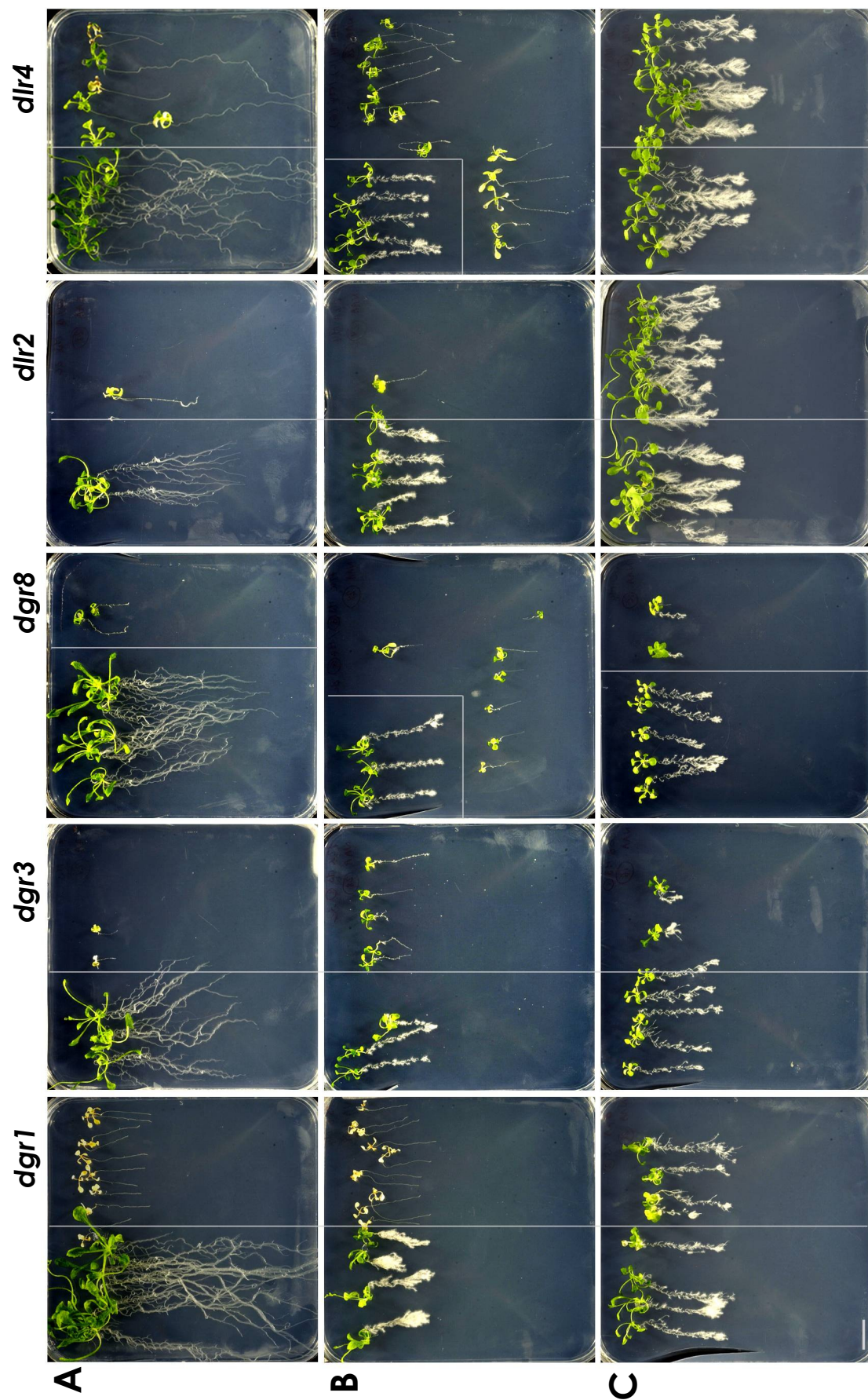


FIGURE 2.3 EFFECT OF EXOGENOUS AUXIN ON THE PHENOTYPE OF *dgr* AND *dlr* MUTANTS. Plants were grown for seven days at 18°C (permissive temperature) in the absence of exogenous auxin. Subsequently, they were transferred to the restrictive temperature (28°C) (A), or they were first transferred to fresh medium supplemented with 10^{-6} NAA and then transferred to 28°C (B) or they were first transferred to fresh medium supplemented with 10^{-6} NAA and further grown at 18°C (C). Pictures were taken four days later. Dotted lines separate wild type (left) from mutant (right) plants. Bar = 1 cm



FIGURE 2.4 EFFECT OF TEMPERATURE ON GROWTH AND DEVELOPMENT OF DIFFERENT TEMPERATURE-SENSITIVE MUTANTS.

(RESULTS OBTAINED BY ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SARTILMAN) Plants were germinated and grown at the permissive temperature of 18°C until the first flower buds appeared. Then they were transferred to the restrictive temperature of 28°C. Pictures were taken two weeks after transfer to the restrictive temperature.

Pictures (A) through (H) (left: restrictive temperature; right: permissive temperature), and (M) illustrate culture on soil; pictures (I) through (L) illustrate hydroponic culture.

(A) WT; (B) *dgr1*; (C) *dgr6*; (D) *dgr2*; (E) *dgr7*; (F) *dlr2*; (G) and (H) necrotic *dgr1* plants; (I) flowers of *dgr3* at the permissive temperature, and (J) at the restrictive temperature; (K) hydroponic cultivation of *dgr6* at the restrictive temperature, and (L) at the permissive temperature; (M) flowers of *dgr6* at the restrictive temperature.



FIGURE 2.5 ROOT SYSTEMS OF WT (LEFT) AND *dgr6* (RIGHT) PLANTS (RESULTS OBTAINED BY ANTOINE ADAMANTIDIS, PIERRE LEJEUNE AND SANDRA ORMENÈSE, UNIVERSITY OF LIÈGE, SART-TILMAN). Plants were cultivated in a hydroponic system for three weeks at the permissive temperature followed by two weeks at the restrictive temperature.

2.3.5 FURTHER PHENOTYPIC ANALYSIS OF THE MUTANTS *dgr1* AND *dgr2*

In the next experiments as well as in the following chapters three and four we have focussed our efforts on the phenotypic and molecular characterization of the mutants *dgr1* and *dgr2*. We selected these two mutants, defective in overall growth, since the main goal of my research was to isolate genes essential for the growth and development of the whole plant. Furthermore, *dgr1* was selected for its irreversible phenotype, whereas *dgr2* was chosen as a representative for the subclass of mutants with a reversible phenotype.

TABLE 2.1 SUMMARY OF PHENOTYPIC ANALYSIS OF THE 12 *ts* ARABIDOPSIS MUTANTS

Class	Name	Plant phenotype						Effect of exogenous auxin		Callus initiation		In vitro growth analysis of the primary root	
		In vitro		On soil		In hydroponic system		18°C	28°C	18°C	28°C	18°C	28°C
		18°C	28°C	18°C	28°C	18°C	28°C						
Class 1	<i>dgr1</i> ²						•Deficiencies in lateral root formation •Overall growth defects	Exogenous auxin does not induce a changed phenotype in comparison to the WT	•Temperature-sensitive phenotypes not influenced by exogenous auxin •No difference between the mutant phenotypes in the presence or in the absence of exogenous auxin	no	no	•Normal growth of the main root ¹ •Normal overall growth ¹	•Growth decrease and finally growth arrest of the main root ² •Growth of the shoot organs is repressed ³
	<i>dgr2</i>						Flower abortion and sterility						
	<i>dgr3</i> ¹						•Deficiencies in lateral root formation •Overall growth defects						
	<i>dgr4</i>						Severely necrotic, flower abortion, apex does not bolt						
	<i>dgr5</i>		•No lateral roots •Overall growth is inhibited		Inhibition of growth and development ²	•Normal root growth •Normal overall growth							
	<i>dgr6</i>												
	<i>dgr7</i>	•Formation of normal lateral roots					•Deficiencies in lateral root formation •Overall growth defects						
	<i>dgr8</i>	•Normal growth (exceptions ¹)		Normal growth ¹									
Class 2	<i>dlr1</i>		•No lateral roots				•	Temperature-sensitive phenotypes are rescued by exogenous auxin	yes	yes	yes	•	•
	<i>dlr2</i>		•Main root and shoot growth is normal				•					•	•
	<i>dlr3</i>		•Less and smaller lateral roots				•					•	•
	<i>dlr4</i>		•Main root and shoot growth is normal				•					•	•

In contrast to all the other mutants, *dgr3*, *dgr4*, and *dgr8* show a slower growth than the WT at the permissive temperature.

²The phenotype of the mutant *dgr1* is the most strongly affected by high temperature since lethality already occurs after four days at the restrictive temperature

¹Experiment is not applicable to the mutant

³No data available

2.3.5.1 MICROSCOPIC STUDY OF THE PRIMARY ROOT MERISTEM

By means of microscopic visualization of the primary root meristem and cortical cell length measurements going from the root tip to the beginning of the mature root zone, we tried to obtain a better insight in the biological process(es) that are affected in both the mutants *dgr1* and *dgr2* (figure 2.8).

No clear differences neither in root structure nor shape were observed between WT and *dgr1* roots. Only in a few cases we noticed cells that were slightly misshaped. However, this could be due to plasmolysis, caused by the method of slide preparation.

In contrast, the meristematic zone in *dgr2* roots was dramatically reduced at the restrictive temperature. The more distant from the root tip, the more cells resembled WT cortical cells again. Striking is the severely swollen and misshaped nature of the cortical cells in this mutant (figure 2.8).

Cell lengths in cortical cell files were measured going from the quiescent center to the mature root zone. When cell length is plotted in function of root length, it is clear that throughout the root growth zone *dgr1* cells are significantly shorter than WT cells. Furthermore, an increased growth zone (meristem and elongation zone) is observed, which has almost twice the length of the growth zone in the WT or in the *dgr1* mutant root grown under permissive conditions (figure 2.9).

Dgr2 cortical cells are significantly larger in the growth zone all the way to the quiescent center. It is unlikely that such large cells in the vicinity of the quiescent center are still dividing. Therefore, a reduction of the meristem in the roots could be the cause of the root growth arrest in this mutant. Furthermore, the growth zone in the root is only half of the length of the growth zone in WT or *dgr2* mutant roots at the permissive temperature.

2.3.5.2 MENDELIAN SEGREGATION ANALYSIS

Mutant *dgr1* and *dgr2* M₃-plants were back-crossed with WT Col plants at the permissive temperature. Through selfing of the resulting BC₁-plants BC₁F₂ seed batches were obtained for each mutant line. For both mutants the segregation was scored by screening the individuals of the BC₁F₂ generation for defects in lateral root formation at the restrictive temperature. None of the tested *dgr1* and *dgr2* populations showed a segregation ratio different from the expected 3:1 ratio at a 5% significance level (table 2.2).

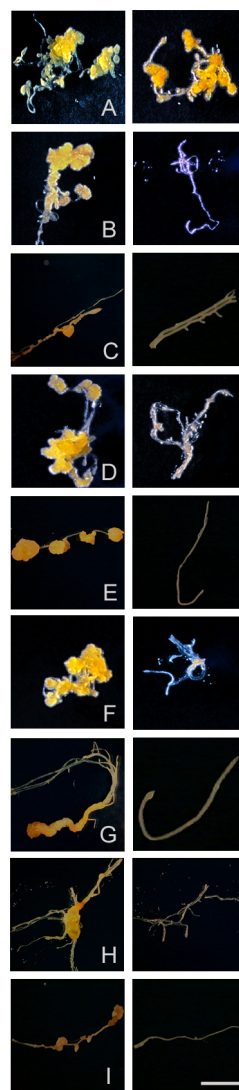
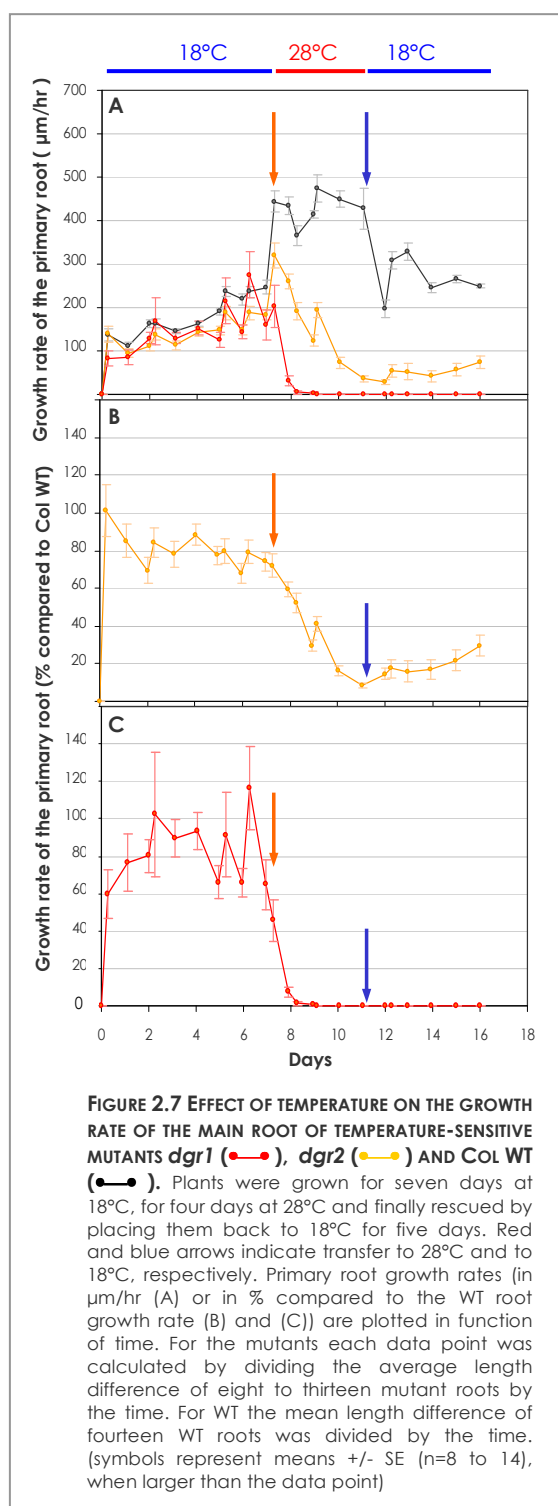


FIGURE 2.6 CALLUS INITIATION FROM ROOT EXPLANTS OF WT (COL) AND THE EIGHT *dgr* MUTANTS. Root explants were transferred to a callus inducing medium and either placed at the permissive temperature of 18°C (left) or the restrictive temperature of 28°C (right). Pictures were taken after 6 weeks of culture. (A) WT; (B) *dgr1*; (C) *dgr3*; (D) *dgr7*; (E) *dgr5*; (F) *dgr6*; (G) *dgr2*; (H) *dgr8*; (I) *dgr4*. Bar=1cm



2.4 DISCUSSION

We isolated twelve novel ts mutants of *Arabidopsis thaliana* that are defective in growth and development. By selecting for deficiencies in lateral root formation, we aimed to isolate these growth mutants in a reasonably fast and straightforward manner. Furthermore, we used exogenous auxin during the entire screening process, since it is expected to be easier to distinguish mutant plants that are not able to form lateral roots from WT plants producing many lateral roots in the presence of auxin.

A general phenotypic analysis allowed to classify the twelve mutants into two categories: (i) a first category consisted of eight mutants (*dgr1* to *dgr8*) that were inhibited in overall plant growth; the observed phenotype was either reversible (*dgr2* to *dgr8*) or irreversible (*dgr1*), which, at the higher temperature, resulted in growth arrest or lethality, respectively; (ii) a second category comprised four mutants (*dlr1* to *dlr4*) that showed specific defects in lateral root formation. Of these, two mutants (*dlr1* and *dlr2*) were not able to form lateral roots at all, whereas the remaining two mutants (*dlr3* and *dlr4*) produced smaller lateral roots compared to WT plants. At a permissive temperature of 18°C the phenotypes of the ts mutants were comparable to those of WT plants.

2.4.1 In *dgr* Mutants Temperature-Sensitivity Is Conferred Throughout Vegetative and Reproductive Development

When grown on soil, *dgr* mutants showed an inhibition of growth and development at the restrictive temperature. For some mutants, this inhibition was quite severe as in *dgr1*, where rapid necrosis and lethality was observed, and in *dgr6*, where growth at

FIGURE 2.8 CELLULAR PHENOTYPE IN THE PRIMARY ROOT OF *dgr1* AND *dgr2*. Microscopic (DIC) pictures of the main root of WT (A) and the temperature-sensitive mutants *dgr1* (B) and *dgr2* (C) are shown. After seven days at the permissive temperature, WT, *dgr1* and *dgr2* plants were transferred to the restrictive temperature for five, two and five days, respectively.

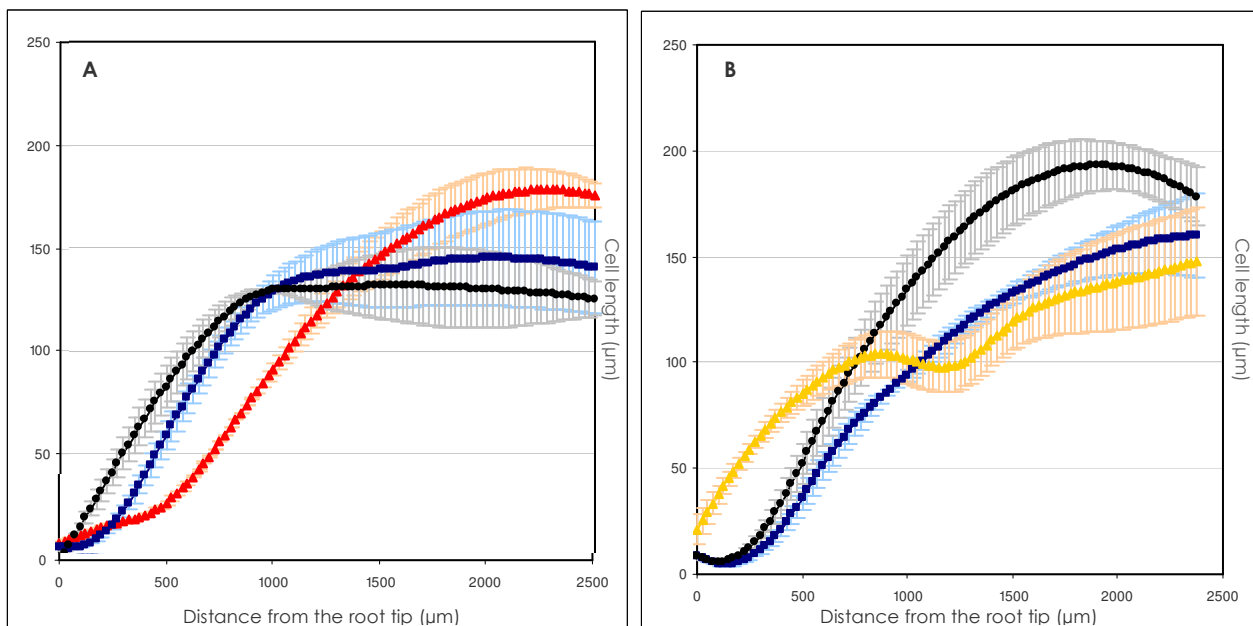
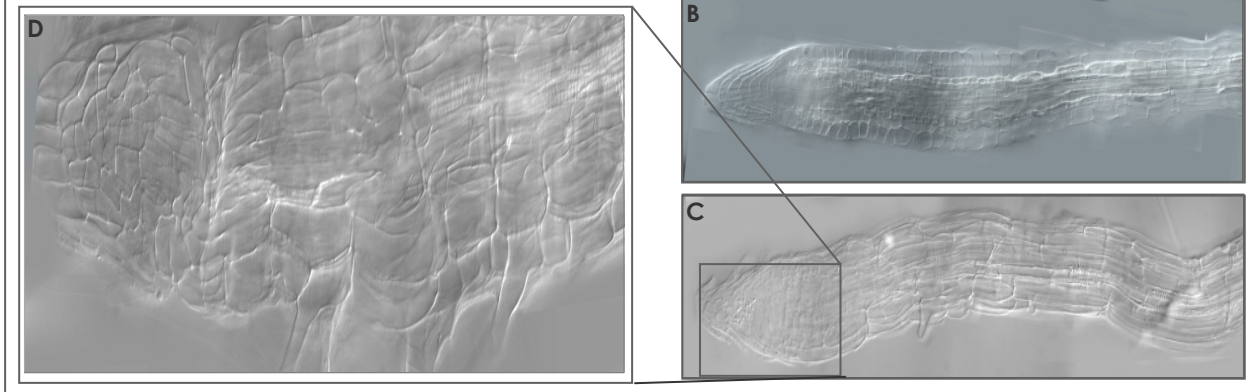


FIGURE 2.9 CORTICAL CELL LENGTHS IN THE PRIMARY ROOTS OF *dgr1* AND *dgr2*. *Dgr1* plants were grown for seven days at 18°C and then for two days at 28°C (A; \blacktriangle — \blacktriangle). *Dgr2* plants were grown for seven days at 18°C and then for five days at 28°C (B; \blacktriangle — \blacktriangle). WT plants were grown simultaneously under the same conditions (A and B; \bullet — \bullet). Control plants (*dgr1* (A) and *dgr2* (B); \blacksquare — \blacksquare) were continuously grown at 18°C. Cell length measurements of the cortical cell file were performed starting from the quiescent center to the mature root zone. (symbols represent means \pm SE (n=5), when larger than the data point)

TABLE 2.2 MENDELIAN SEGREGATION ANALYSIS FOR *ts* MUTANTS *dgr1* AND *dgr2*

CROSS	NUMBER OF PLANTS	EXPERIMENTAL SEGREGATION		THEORETICAL SEGREGATION		X ² -VALUE*
		WTs	Mutants	WTs	Mutants	
<i>ColWT(1)</i> x <i>dgr1(1)</i>	64	48	16	48	16	0.00
<i>ColWT(2)</i> x <i>dgr1(2)</i>	70	55	15	52.5	17.5	0.48
<i>ColWT(3)</i> x <i>dgr1(3)</i>	70	57	13	52.5	17.5	0.48
<i>ColWT(4)</i> x <i>dgr1(4)</i>	40	30	10	30	10	0.00
<i>ColWT(5)</i> x <i>dgr1(5)</i>	40	29	11	30	10	0.13
<i>dgr1(1)</i> x <i>ColWT(1)</i>	70	54	16	52.5	17.5	0.17
<i>dgr1(2)</i> x <i>ColWT(2)</i>	70	54	16	52.5	17.5	0.17
<i>dgr1(3)</i> x <i>ColWT(3)</i>	70	57	13	52.5	17.5	1.54
<i>dgr1(4)</i> x <i>ColWT(4)</i>	70	58	12	52.5	17.5	2.30
<i>dgr1(5)</i> x <i>ColWT(5)</i>	70	55	15	52.5	17.5	0.48
<i>ColWT(1)</i> x <i>dgr2(1)</i>	70	55	15	52.5	17.5	0.48
<i>ColWT(2)</i> x <i>dgr2(2)</i>	70	53	17	52.5	17.5	0.02
<i>ColWT(3)</i> x <i>dgr2(3)</i>	70	54	16	52.5	17.5	0.17
<i>ColWT(4)</i> x <i>dgr2(4)</i>	65	52	13	48.75	16.25	0.87
<i>ColWT(5)</i> x <i>dgr2(5)</i>	70	56	14	52.5	17.5	0.93
<i>dgr2(1)</i> x <i>ColWT(1)</i>	70	53	17	52.5	17.5	0.02
<i>dgr2(2)</i> x <i>ColWT(2)</i>	67	56	11	50.25	16.75	2.63
<i>dgr2(3)</i> x <i>ColWT(3)</i>	70	58	12	52.5	17.5	2.30
<i>dgr2(4)</i> x <i>ColWT(4)</i>	70	53	17	52.5	17.5	0.02
<i>dgr2(5)</i> x <i>ColWT(5)</i>	60	41	19	45	15	1.42

*X_(0.05;1)²=3.84 (for 3:1 ratio)

the higher temperature induced flower abortion and sterility.

Cultivation of the mutants in a hydroponic system furthermore revealed that *dgr* mutants showed disruptions in the formation of secondary roots but also the growth of the root and shoot was affected. The *dgr6* mutant turned necrotic at the apex and did not bolt at 28°C, whereas flowers of *dgr3*, were aborted and sterile.

These results indicate that the *DGR* genes are required for normal root and shoot meristem functioning

throughout vegetative and flower development.

Although *ts* mutations provide an excellent opportunity to define the specific developmental stage(s) at which the protein encoded by the mutated gene acts, in *Arabidopsis* this was only tested in a few studies. The *add* mutants, for instance, that were isolated as nine-day old seedlings with *ts* disruptions in apical meristem function, proved to confer temperature-sensitivity throughout the vegetative stage (Pickett et al., 1996). Older *add1* plants were arrested in

development when shifted from low to high temperature, which indicated that *ADD1* was required not only for the organization or initiation of the shoot apical meristem, but also for normal meristem function throughout the *Arabidopsis* life cycle. A similar result was seen with *add2* plants.

Our *dgr* mutants showed temperature-sensitivity throughout the vegetative and the reproductive developmental stage, suggesting that the mutated genes play a vital role during the entire plant life cycle. Such genes may be involved in processes that are required to be permanently active, such as house keeping genes or genes involved in cell proliferation processes.

2.4.2 *Dgr* Mutations Affect Processes Essential for Overall Plant Growth

At the restrictive temperature, the overall growth of the plant is inhibited in all *dgr* mutants, implying that the mutations affect genes involved in fundamental processes of growth and development, possibly the same in shoot and root. The temperature effects are reversible for seven mutants when the shift to the restrictive temperature is less than four days. Only *dgr1* shows an irreversible mutant phenotype and becomes necrotic and dies after four days at the restrictive temperature.

Several mutants with an altered morphology of the whole plant have been described in literature. In the *glabra2* (*gl2*) mutant, epidermal patterning is altered both in the root and in the shoot suggesting that *GL2* controls epidermal cell fate in the root and the shoot (Hung et al., 1998). Another set of genes appears to perform the same function in root and shoot in ground tissue patterning: *SHORT-ROOT* (*SHR*) and *SCARECROW* (*SCR*) are required for a normal endodermis (also known as 'starch-sheath') in the stem, hypocotyls and the root (Fukaki et al., 1998). These studies indicate that many

genes involved in fundamental cellular processes are simultaneously active in several organs of the plant or even throughout the entire plant. However, using classical mutants, only those genes of which the function is not essential to plant survival, will be picked up.

Therefore, the *dgr* ts mutants, showing an overall growth arrest at the restrictive temperature, and in particular the conditional lethal mutant *dgr1*, are expected to harbour mutations in genes essential for plant life that are simultaneously active in all organs of the plant. Identifying such genes that function in the fundamental set of processes that are required to produce a viable plant with a normal phenotype, will contribute a lot to the understanding of the basic mechanisms of plant growth and development.

2.4.3 *Dlr* Mutations Specifically Disrupt Lateral Root Formation

Lateral root formation is specifically affected in the *dlr* mutants. The mutated genes of *dlr1* and *dlr2* seem to regulate the initiation steps of the lateral root formation while the mutated genes of *dlr3* and *dlr4* probably control the subsequent stages during which meristem organization takes place within the primordium. Three *aberrant lateral root formation* (*alf*) mutants with altered position, development or number of lateral roots were isolated by Celenza et al. (1995). The *alf1-1* mutant forms an increased number of lateral roots. The *alf3-1* mutant is defective in the development of lateral root primordia into mature lateral roots such as the phenotype of mutants *dlr3* and *dlr4*. On the other hand, the phenotype of the *alf4-1* mutant is comparable to that of mutants *dlr1* and *dlr2* in which no lateral roots can be formed. The *alf3-1* mutant can be rescued by auxin (IAA), whereas this is not possible for the *alf4-1* mutant. Since the *alf4* mutant shows only defects in lateral root formation

and not in primary root formation, it is suggested that certain genes exist, which specifically function in the process of lateral root formation.

De Smet et al. (PhD thesis) used a genome-wide transcript profiling approach to identify genes, specifically expressed during lateral root initiation. Based on their transcript profiling, they recovered and analyzed mutants for eighteen genes expressed during lateral root initiation. Mutations in three genes (at2g33830; dormancy/auxin associated family; at4g28640; IAA11 and at2g28620; kinesin motor protein-related) strongly reduced the number of lateral roots per centimeter with more than 40%. However, these mutants also showed a slower growth of the main root.

Since only very few mutants are described that are specifically affected in lateral root formation and show normal growth of all other plant organs, the further characterization of the *dlr* mutants would be useful to identify the genes with such specific functions.

For the research described in the next chapters, we have focussed our efforts on the molecular characterization of the mutants *dgr1* and *dgr2*. We selected these two mutants, defective in overall growth, since our main goal was to isolate genes essential for the growth and development of the whole plant. Furthermore, *dgr1* was selected for its irreversible phenotype, whereas *dgr2* was chosen as a representative for the subclass of mutants with a reversible phenotype.

2.5 CONCLUSION

In a first aspect, our study demonstrated that *ts* mutants provide an extremely powerful tool for studying gene function *in vivo*. They represent a reversible mechanism to lower the level of a specific gene product at any stage in the growth of the organism simply by changing the temperature of growth.

In a second aspect, we have isolated twelve interesting mutants, each potentially corresponding to a novel (essential) gene. Further phenotypic and molecular genetic characterization of the *ts* growth mutants, described in this chapter, is expected to provide various information, which will be useful for the further understanding of the biological mechanisms underlying plant growth and development.

ACKNOWLEDGEMENTS

I would like to thank Dr. Sandra Ormenèse, Prof. Dr. Annie Jacqmard and Dr. Pierre Lejeune for the large amount of work they invested in the mutant screening procedure and the initial phenotypic characterization of the ts mutants.

CHAPTER THREE

MAP-BASED CLONING OF THE TEMPERATURE-SENSITIVE MUTATIONS *dgr1* AND *dgr2* IN *ARABIDOPSIS*

In this study, we describe the map-based cloning of two temperature-sensitive mutations, *dgr1* (defect in growth 1) and *dgr2* (defect in growth 2) that we previously identified by screening an EMS mutagenized population of *Arabidopsis*. Plants harbouring the *dgr1* or *dgr2* mutations are rapidly arrested in growth when transferred to the restrictive temperature. Furthermore, *dgr1* exhibits irreversible growth defects, eventually resulting in lethality after four days at the restrictive temperature.

Using an AFLP-based genome-wide mapping strategy, we mapped both mutations to a chromosomal segment of the *Arabidopsis* genome. This approach proved to be very useful in the initial stage of the map-based cloning process, when no knowledge about the location of the gene of interest was yet available.

For the subsequent fine-mapping of *dgr1*, we used individual PCR-based InDel markers from the *Arabidopsis* Cereon collection. As a result, we succeeded in cloning the *DGR1* gene or at2g45790, and revealed that it encodes an eukaryotic phosphomannomutase. This enzyme was never functionally characterized before in *Arabidopsis* but has been known to play a crucial role in several fundamental cellular processes, including cell wall biosynthesis, glycoprotein synthesis and plant ascorbic acid biosynthesis in other species. Through our work on the *dgr1* mutant, we have cloned and characterized a novel essential plant gene and also provided a new and useful tool for further dissection of the several (potentially unknown) pathways involving phosphomannomutase in *Arabidopsis*.

The *DGR2* gene was mapped to a chromosomal region of 3.5Mb on the top of chromosome II, consisting of NOR2 and the upper telomere. This 3.5Mb region contains only a very small sequenced 38.5 kb fragment preceding the NOR region, hampering further fine-mapping of the *DGR2* gene. Despite the fact that NOR2 and the telomeric region largely consist of non-protein coding genes, they fulfil important functions in plant growth and development, which will be discussed in detail in this chapter.

3.1 INTRODUCTION

Now that the genome of *Arabidopsis* has been sequenced, this cruciferous plant has emerged as the preferred reference organism for understanding the function of unknown genes in higher plants. The most common method for elucidating the function of unknown genes is to use various mutagenesis procedures, such as insertional mutagenesis (Feldmann, 1991; Koncz et al., 1992; Bancroft and Dean, 1993; Aarts et al., 1995), gene silencing (Baulcombe, 1996; Kooter and Mol, 1993), and physical or chemical mutagenesis (Redei and Koncz, 1992). Gene disruption by insertional mutagenesis can be accomplished by either transposon mutagenesis (Bancroft and Dean, 1993; Aarts et al., 1995) or *Agrobacterium*-mediated T-DNA transformation (Feldmann, 1991; Koncz et al., 1992). The advantage of insertional mutagenesis is that the gene affected can be identified easily. However, with few exceptions, such as the reported overexpression of *PIF3* (for phytochrome-interacting factor) after T-DNA insertion in its promoter region, insertional mutagenesis lacks the versatility to create partial loss-of-function or gain-of-function alleles, which are often very informative in elucidating the biological function of genes. In contrast, chemical or physical mutagenesis can generate simple sequence mutations, such as point mutations or small deletions and insertions, and thus can be used to dissect functional domains of proteins by producing a variety of different phenotypes associated with a specific genetic locus. The major drawback in generating mutations by chemical or physical means is the ability to identify the locus responsible for the mutant phenotype because locus identification, unless limited to a few candidate genes (McCallum et al., 2000), entails map-based cloning, a process that until recently has been considered time-consuming and expensive. However, with the public accessibility of the complete

and annotated sequence of the *Arabidopsis* genome, this situation has radically changed. Chromosome walks, formerly the most tedious and technically problematic steps in the process, are a thing of the past. Nowadays, molecular mapping in *Arabidopsis* is vastly enhanced by systematically exploiting the available sequence information.

Sequence variation between individuals includes single nucleotide polymorphisms (SNPs) and deletions/insertions (InDels). These make up the backbone for the mapping and positional cloning of genes responsible for phenotypic traits. By tracking DNA sequence variants that cosegregate with heritable traits, the genes giving rise to these traits can be localized to specific chromosomal locations. The majority of known DNA sequence variants in use over the past 20 years have been polymorphisms affecting the recognition sequence of a restriction endonuclease, commonly referred to as restriction fragment length polymorphisms (RFLPs) (Botstein et al., 1980). With the development of the polymerase chain reaction (PCR), novel DNA fingerprinting methods were developed, such as CAPS (cleaved amplified polymorphic sequences; Weining and Langridge, 1991; Konieczny and Ausubel, 1993) and AFLP (Vos et al., 1995). The CAPS method uses amplified DNA fragments that are digested with a restriction endonuclease to display any polymorphic restriction sites, and the products are analyzed by agarose gel electrophoresis. Using the AFLP method, sets of restriction fragments are visualized by PCR without knowledge of nucleotide sequence.

In this study, we used an AFLP-based genome-wide mapping strategy (Peters et al., 2001; Peters et al., 2004) to map the two temperature-sensitive (ts) mutations in *Arabidopsis*, called *dgr1* and *dgr2*. This genome-wide mapping strategy proved to be very helpful at the initial steps of the map-based cloning process, when no knowledge about the location of the gene of interest is available. Briefly, for

each mutation we analyzed approximately 120 F₂ segregants with eight primer combinations (PCs), providing a well-dispersed grid of 85 AFLP markers covering the genome. Segregation analysis established linkage to a chromosomal region of less than 5 Mb. For *dgr1* and *dgr2*, segregation analysis of additional AFLP markers in the target region resulted in a chromosomal segment of less than 700 kb and 3.5 MB, respectively.

For the subsequent fine-mapping of the mutated gene the advantage of AFLP (i.e. many physically dispersed markers per PC) is lost and therefore further segregation analyses were based on individual, robust, PCR-based markers (InDels) from the Cereon collection (<http://www.arabidopsis.org/Cereon/index.html>).

Using the above-described approach, we were able to identify the *DGR1* gene as a phosphomannomutase, an essential enzyme functioning in several biological processes, such as glycosylation and vitamin C (ascorbic acid) production in plants. Moreover, we mapped the *DGR2* gene to a chromosomal region of 3.5Mb on the top of chromosome II. Unfortunately, sequence information is unavailable for this chromosomal segment, which mainly consists of the highly repetitive nucleolar organizing region (NOR) and the upper telomere of chromosome II.

3.2 EXPERIMENTAL PROCEDURES

3.2.1 MUTAGENESIS

The plant material used in the present study was derived from the Columbia-0 strain (Col-0) of *Arabidopsis thaliana* (L.) Heynh that is referred to as the wild type (WT) in this paper. Chemical mutagenesis was performed with ethyl-methane sulfonate (EMS) (15µM) for 24

hours at 24°C on a population of 200,000 seeds (M₁ generation). The M₂ progeny was screened for ts mutations. Subsequently, heritability of the observed phenotype was tested with M₃ plants produced by self-fertilization of M₂ plants of putative mutant lines.

3.2.2 GROWTH CONDITIONS

Seeds were sterilized by a ten-minute incubation with a 1.5% sodiumhypochlorite solution (Merck, Germany). Sterile seeds were sown on MS germination medium (Murashige and Skoog, 1962) in square Petri dishes (120mm x 120mm, Greiner bio-one, Frickenhausen, Germany). After sowing, Petri dishes were placed vertically in a growth chamber (temperature specified below) under continuous light provided by V.H.O. Sylvania fluorescent white tubes at an irradiance of 50 µmol PAR photons m⁻².s⁻¹ (Osram, Milan, Italy) or by Lumilux Cool White light tubes (Osram, Milan, Italy) at an irradiance of 50 µmol PAR photons m⁻².s⁻¹. ts phenotypes were identified *in vitro* by growing plants for seven days at a permissive temperature of 18°C and transferring them subsequently to a restrictive temperature of 28°C (further experimental details are described in chapter 2).

3.2.3 PLANT MATERIAL

EMS is a mutagen that typically induces hundreds of point mutations per genome (for more details see chapter one: § 1.3.2.1 and figure 1.11). To avoid phenotypic interference of side mutations (different from the mutation in the gene of interest), it is usually recommended to perform at least one back-cross between the mutant line of interest and the wild type (of the same accession), before initiating the different successive crosses

and selfings needed for generation of a mapping population (figure 3.1).

For both *dgr1* and *dgr2* one back-cross cycle was performed. Mutant M_3 -plants (grown on soil) were back-crossed with Col-0 plants (grown on soil) at the permissive temperature (18°C). This was done under a binocular (Leica MZ 16F, Germany) using tweezers to transfer pollen from the WT Col-0 plant to the stamen of the M_3 -plant. Through selfing of the resulting BC_1 -plants, BC_1F_2 seed batches were obtained. The segregation of the BC_1F_2 -families was tested for a 3:1 ratio of WT to mutant phenotypes. To this end, at least 50 plants were grown for seven days *in vitro* on MS medium at the permissive temperature and subsequently transferred to the restrictive temperature (28°C). Subsequently, BC_1F_2 plants showing a mutant phenotype at the restrictive temperature were rescued by transferring them to soil and back to the permissive temperature. Self-pollination of these mutant plants led to the production of

dgr1 and *dgr2* lines, for which the genetic background was purified for 50 percent. The homogeneity of the resulting mutant lines was tested by growing at least 50 plants for seven days at the permissive temperature, subsequently transferring these plants to the restrictive temperature, and checking the consistency of the mutant phenotype.

Next, to generate mapping populations for both *dgr1* and *dgr2*, back-crossed M_3 mutant lines were reciprocally crossed with Ler WT (figure 3.1). Heterozygosity of the resulting F_1 plants was checked using the microsatellite marker *nga8*, which is polymorphic between Col and Ler. F_2 seeds were harvested from those F_1 plants that were confirmed to be heterozygous. The resulting segregating *dgr1* and *dgr2* F_2 families served as mapping populations, used throughout the subsequent map-based cloning process. Each F_2 plant was phenotyped prior to harvesting and being frozen in liquid nitrogen.

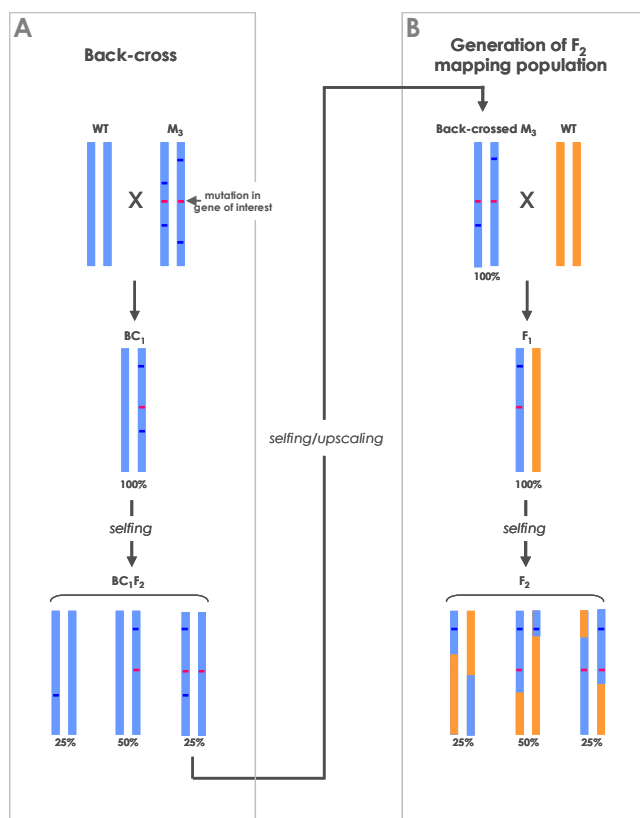


FIGURE 3.1 SCHEME OF THE DIFFERENT SUCCESSIVE CROSSES AND SELFINGS PERFORMED FOR BOTH *dgr1* AND *dgr2* PRIOR TO THE MAP-BASED CLONING PROCESS.

Firstly (A), one back-cross cycle was performed for both *dgr1* and *dgr2* to avoid phenotypic interference of side mutations (■), which are typically induced when using EMS as a mutagen (for more details see chapter one: §1.3.2.1 and figure 1.11). Mutant M_3 plants were back-crossed with WT plants of the same accession (e.g. Col: blue; pairs of vertical lines represent homologous chromosomes). After selfing of the resulting BC_1 -plants BC_1F_2 seed batches were obtained. A 3:1 segregation ratio of WT to mutant phenotypes was observed for both *dgr1* and *dgr2* BC_1F_2 -families. Selfing of mutant BC_1F_2 -plants led to the production of back-crossed *dgr1* and *dgr2* lines, for which the genetic background was purified for 50 percent.

Next, to generate F_2 mapping populations for *dgr1* and *dgr2* (B), back-crossed M_3 lines were crossed with the WT of a different accession (e.g. Ler: orange). Through selfing of the heterozygous F_1 plants resulting from this cross F_2 seeds were obtained. The resulting F_2 -families segregating for either *dgr1* or *dgr2*, served as mapping populations used throughout the subsequent map-based cloning process.

3.2.4 DNA EXTRACTION

The method for DNA preparation described by Murray and Thompson (1980) (cTAB method) was modified as described below.

Individual F₂ plants were ground in separate tubes (Eppendorf; Germany). Immediately after homogenization, the tissue powder was added to 400 µL cTAB-buffer (200mM Tris-HCl pH 7.5; 2M NaCl; 50 mM EDTA; 2% (m/v) cTAB) followed by a 60 minute incubation at 65°C. Samples were subsequently put on ice for 15 minutes. 250 µL chloroform/iso-amylalcohol (24:1) was added and samples were mixed and centrifuged at full speed (13,000 rpm) for ten minutes at room temperature. The upper DNA containing aqueous phase was then transferred to a fresh eppendorf tube containing 200µL isopropanol. Samples were mixed thoroughly and centrifuged for 15 minutes at full speed (13,000 rpm). Hereafter, the supernatants was discarded and the pellet was rinsed with ethanol (70%), air-dried and suspended in 50µL of TE (0.1 mM EDTA).

3.2.5 AFLP ANALYSIS

AFLP template preparation was carried out according to Vos et al. (1995) starting from 100 ng DNA. The restriction enzymes used were *SacI* and *MseI*. For the pre-amplifications, a selective *MseI* primer was combined with a *SacI* primer, each containing a single selective nucleotide at the 3' end. PCR conditions were as described by Vos et al. (1995). The amplification products were diluted 20-fold and 5 µL aliquots were used for final selective amplifications with *SacI*+2/*MseI*+2 PCs and *Taq* polymerase (Roche). The selective *SacI*+2 primers were radioactively phosphorylated using [γ -³²P]ATP (specific activity >92 TBq mmol⁻¹; Amersham, Buckinghamshire, UK) and the DNA fragments were separated in a Sequi-Gen GT sequencing cell (Bio-Rad,

Hercules, CA). As a size marker, the SequaMark 10 base-ladder (Research Genetics, Huntsville, AL) was used. After electrophoresis, gels were dried on a Heto dry GD-1 slab gel dryer (Heto Lab Equipment, Allerød, Denmark) and visualized using a PhosphorImager 445 SI (Molecular Dynamics, Sunnyvale, CA). Markers were scored using AFLP Quantar Pro™ software (Keygene Products B.V., Wageningen, The Netherlands). AFLP is a registered trade mark of Keygene N.V.

3.2.6 WHOLE-GENOME MAPPING USING A PHYSICALLY ANCHORED ARABIDOPSIS AFLP MAP

In this study, the physically anchored *Arabidopsis* AFLP map published by Peters et al. (2001) was used. This map was developed based on the genome sequence of *Arabidopsis thaliana* (Col) (*Arabidopsis* Genome Initiative, 2000) and the AFLP high-throughput marker technology (Vos et al., 1995). Briefly, for the construction of the physically anchored map, experimental and *in silico* AFLP were combined, using *SacI* and *MseI* as restriction enzymes. As a result, a collection of markers present in Col and absent in Ler (referred to as Col markers), and a database of *in silico* AFLP fragments, based on the Col-0 genome sequence, were obtained. By combining the size of an experimental AFLP fragment with the selectivity of the AFLP primers and the rough genetic position of the fragment, determined by the segregation in a limited set of ColxLer RILs (Lister and Dean, 1993), a specific experimentally observed marker could be linked to a specific *in silico* AFLP fragment in the database. Using this approach, 1266 Col markers were unambiguously positioned on the genome sequence of *Arabidopsis thaliana*.

Based on the above described map, Peters et al. (2004) defined a set of 8 PCs, generating 85 AFLP markers that are well dispersed over the genome. This

limited set of PCs allowed us to detect the chromosomal segment containing the gene of interest (table 3.1; figure 3.2). Additional AFLP markers positioned in the identified interval were used to further delineate the region of interest. At this stage the benefits of a detailed mapping position must be balanced against the work involved in the AFLP-characterization of a larger number of F_2 individuals. Usually, in the case of *Arabidopsis*, analyzing about 400-450 F_2 individuals by the currently available set of 1266 physically mapped AFLP markers results in a region of 400-700 kb containing the gene of interest. Details on the position of the AFLP markers in the *Arabidopsis* genome can be found on the TAIR website (<http://www.arabidopsis.org/>).

3.2.7 FINE-MAPPING USING INSERTION/DELETION (INDEL) POLYMORPHISMS

For fine-mapping, a large population of F_2 -individuals (6,000-10,000) was screened with markers flanking the region containing the gene of interest. We used PCR-based InDel markers from the Cereon collection of InDel polymorphisms (Cereon Genomics, Cambridge, MA, USA; <http://www.Arabidopsis.org/Cereon/index.html>).

InDel polymorphisms were identified, and primers developed (using the software Primer3™ version 0.2; Whitehead Institute, Cambridge, UK and Howard Hughes Medical Institute, Chevy Chase, Maryland, USA), to score the flanking markers in a co-dominant way (figure 3.3). The conditions used for PCR amplification were: 2 minutes at 94°C; 10 cycles of 15 seconds at 94°C, 30 seconds at 63°C, 30 seconds at 72°C with a touch-down of 1°C per cycle; 35 cycles of 15 seconds at 94°C, 30 seconds at 53°C, 30 seconds at 72°C; 2 minutes at 72°C. Each time, two reactions were run on a single polyacrylamide gel. Depending on the size of the PCR-products, samples were

loaded simultaneously or separated in time.

Since both mutations are recessive, homozygous and heterozygous WT plants could not be distinguished phenotypically in an F_2 population. One way to deduce the zygosity of the WT F_2 individuals is to generate F_3 progeny by selfing each WT F_2 individually and check the segregation. This is, however, a time-consuming process. We used an alternative and much faster approach by performing the InDel linkage analyses only on mutant F_2 individuals. This way, more F_2 plants have to be grown initially to obtain a sufficient number of mutant individuals, but, as opposed to the first option, it saves the time needed to complete an entire *Arabidopsis* life cycle.

The identified recombinants can subsequently be used to further delineate the area containing the gene of interest with additional markers from the available Cereon InDel and SNP collection, positioned within the region of interest.

3.2.8 SEQUENCE ANALYSIS

Genes (exons and introns) were sequenced as fragments of approximately 500 bp with a minimal overlap of 100 bp.

PCR primers were designed using the Primer3™ software (version 0.2; Whitehead Institute, Cambridge, UK and Howard Hughes Medical Institute, Chevy Chase, Maryland, USA). For amplification of the 500 bp fragments, a PCR was performed with 50 ng genomic DNA template, a final concentration of 0.5 μ M for each primer and using Silverstar DNA polymerase (Eurogentec, Liège). The conditions used for the PCR amplification were: 2 minutes at 95°C; 35 cycles of 30 seconds at 95°C, 30 seconds at average T_m , 2 minutes at 72°C; 5 minutes at 72°C; 4°C. PCR products were purified using ExoSAP-IT™ (USB, Cleveland, Ohio, USA). Purified PCR products (ideal final concentration 10ng/100bp) were deposited on the wells of a semi-coated

PCR plate together with 50µM of primer. The plate was subsequently centrifuged at 4,000 rpm in a 96-well plate centrifuge (Eppendorf, 5810R) and dried at 65°C for 0.5 to 2 hours. Sequencing mix and buffer were added and the plate was spun briefly at 4,000 rpm (Eppendorf, 5810R). A second PCR was then performed under the following conditions: 2 minutes at 95°C; 40 cycles of 10 seconds at 95°C, 10 seconds at 48°C, 2 minutes at 60°C; 4°C. Samples were sequenced with an automated sequencer (ABI3700, ABI377; Applied Biosystems Inc.).

3.3 RESULTS

3.3.1 MAP-BASED CLONING OF *dgr1*

The *dgr1* mutant was identified by screening an EMS-mutagenized population of *Arabidopsis thaliana* (ecotype Col-0) for ts growth defects at a restrictive temperature of 28°C. At the permissive temperature of 18°C *dgr1* plants showed a WT phenotype.

The isolation and phenotypic characterization of the mutant have been described in detail in the previous chapter. After back-crossing *dgr1* M₃ plants with Col-0 WT plants, a segregation analysis of the BC₁F₂ families revealed the recessive nature of the mutation (chapter two: table 2.2).

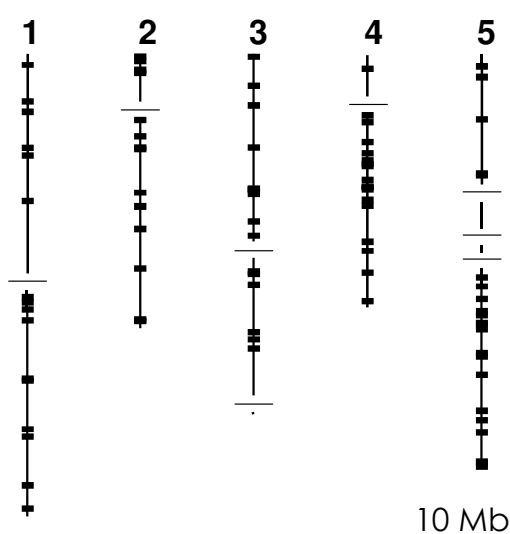


FIGURE 3.2 DISTRIBUTION OF THE 85 COL/LER AFLP MARKERS OVER THE FIVE CHROMOSOMES OF ARABIDOPSIS (PETERS ET AL., 2004). Markers were generated by eight AFLP PCs. On average, one Col/Ler AFLP marker is present every 1.3 Mb (6cM). Only 10% of the markers is separated by more than 3.5 Mb, whereas the maximum gap is about 5.6 Mb. Each bold horizontal line represents a marker, whereas a thin horizontal line disrupting the chromosome, stands for a sequence gap.

F₂ mapping populations were obtained by performing reciprocal crosses of *dgr1* mutant plants (Col-0) with Landsberg erecta (Ler) WT plants (figure 3.1). Each F₁ cross was checked for heterozygosity with the SSLP marker 'nga8' polymorphic between Col and Ler (figure 3.4). F₂ seeds were subsequently harvested from those F₁ plants confirmed to be heterozygous. The segregating F₂ populations were used for the subsequent map-based cloning process (figure 3.5 and 3.6). To this end, the phenotype of each individual F₂

TABLE 3.1 SET OF EIGHT AFLP PCs USED TO DETECT LINKAGE BETWEEN 85 COL/LER AFLP MARKERS AND THE LOCUS OF INTEREST.

MARKER CODE ¹	SELECTIVE NUCLEOTIDES		NUMBER OF AFLP MARKERS
	SACI + 2	MSEI + 2	
SM8	AA	CT	9
SM57	AT	GA	13
SM61	AT	TA	9
SM205	TA	TA	12
SM229	TG	CA	14
SM233	TG	GA	10
SM236	TG	GT	10
SM240	TG	TT	8

¹ according to Peters et al. (2001)

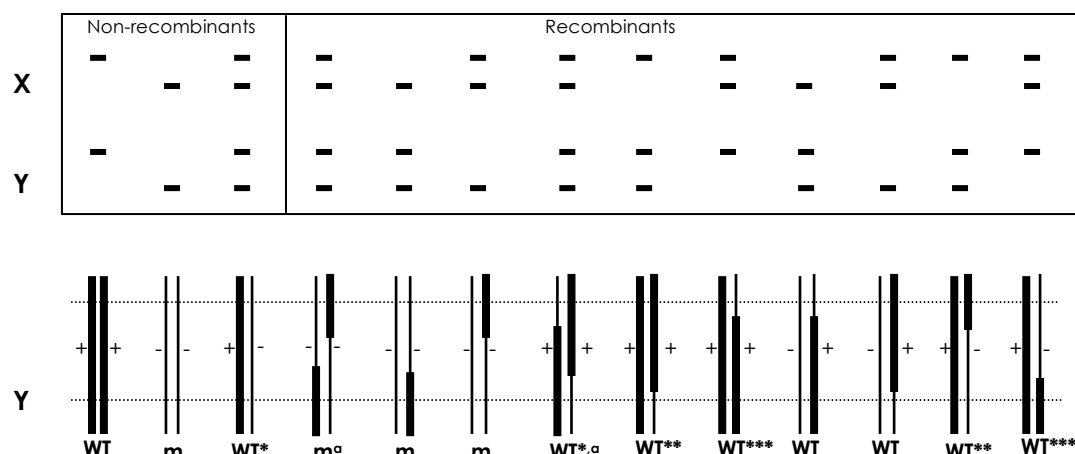


FIGURE 3.3 SCREENING OF A LARGE F₂ POPULATION WITH TWO FLANKING INDEL MARKERS (AFTER PETERS ET AL., 2004). Two markers, X and Y, represent InDel markers flanking the genomic region harbouring the mutation (+: WT allele; -: mutant allele). The F₂ individuals that harbour a recombination event in the region flanked by X and Y can easily be scored, because in all cases but one they produce three PCR fragments instead of two (homozygous F₂ non-recombinants) or four (heterozygous F₂ non-recombinants). Further delineation of the region of interest is an iterative process, using markers localized more inward. The F₂ genotype and phenotype belonging to non-recombinants and recombinants are represented in the bottom panel.

^aIn an exceptional event, four fragments can also be obtained when dealing with a homozygous F₂ recombinant individual.

*, **, *** When dealing with a recessive mutation, homozygous and heterozygous WT plants cannot be distinguished phenotypically in an F₂ population. Furthermore, although different in genotype at the mutant locus, homozygous and heterozygous WT plants may be identical in flanking marker genotypes (compare two WT*, two WT**, or two WT***).

plant was determined quantitatively by measuring the growth rate of the main root as described in paragraph 2.3.4 of the previous chapter. This phenotypic analysis strategy allowed to efficiently distinguish mutant plants from WT plants that showed coincidental growth deficiencies (figure 3.5).

3.3.1.1 WHOLE-GENOME MAPPING OF *dgr1*

In a first step of the genome-wide mapping of *dgr1* a total of 126 F₂ plants, of which 113 WTs and 13 mutants, together with the parental lines Ler and Col, were analyzed using the set of eight PCs outlined in table 3.1. Segregation of the 85 AFLP markers showed clear/complete linkage to a region on chromosome II (figure 3.8). Due to the distances between the AFLP markers resulting from the eight selected PCs (table 3.1; figure 3.2), the genome-wide mapping procedure is expected to determine linkage of the mutant locus to a specific chromosomal region of less than 5.6 Mb, provided that

the required recombinants are present. This is indeed the case for the region containing the *dgr1* mutation, which was delimited to 4.7 Mb, an interval delineated by the upper flanking marker SM240_448,2 (figure 3.6) and the lower end of chromosome II.

In a next step, 13 additional PCs, amplifying AFLP markers in the vicinity and within the defined region of interest (lower half of chromosome II), were used for further linkage analysis performed on the same 126 F₂ individuals (table 3.2). This way, the region containing the *dgr1* locus could be narrowed down to 1.3 Mb with SM26_495,4 as the upper flanking AFLP marker (figure 3.8). Using another eight additional AFLP markers, we did not detect any additional recombination in the 1.3 Mb region. Therefore, as a final step of the AFLP analysis, a larger number of F₂ plants, consisting of 138 WTs and 46 mutants, was grown and analyzed with the remaining AFLP markers localized within the defined 1.3Mb interval. This allowed to further map the *DGR1* gene to a 673 kb region delineated by the upper flanking marker SM120_153,8 and the

lower flanking marker SM190_142,6 (figure 3.9). For the two remaining AFLP markers SM61_107,7 and SM117_193,1, which are localized within the 673 kb region, complete linkage was detected. At this stage, the benefits of a detailed map position had to be balanced against the work involved in the AFLP characterization of an even larger F_2 population. Therefore, fine-mapping of the *dgr1* locus using individual robust PCR markers was initiated.

3.3.1.2 FINE-MAPPING OF *dgr1*

For the fine-mapping of the *dgr1* locus, we grew an F_2 population of 9,900 plants. The phenotype of each individual F_2 plant was determined quantitatively by measuring the growth rate of the main root as described in §2.3.4 of the previous chapter. Since the *DGR1* gene harbours a recessive mutation, homozygous and heterozygous WT plants could not be distinguished solely based on their phenotypes (figure 3.3).

Therefore, to avoid the time-consuming step of selfing the relevant WT recombinants and growing the F_3 progeny to determine their genotype, we only used the mutant F_2 individuals for further analysis. 1,317 F_2 mutants (13.3%) were identified, harvested and submitted to linkage analyses using InDel markers

developed from the Cereon InDel collection (table 3.3).

Initially, two InDel polymorphisms (CER452389 and CER449256) flanking the previously defined 673 kb interval were identified and used to screen the F_2 mutant population. Fourty recombinants were identified and subsequently analyzed with six InDel markers localized in the region delineated by the two original flanking InDel markers. As a result, the interval could be narrowed down to 224 kb with CER452385 and CER452346 as the upper and the lower flanking marker, respectively. The remaining nine recombinants were screened with 11 additional InDel markers localized within the 224 kb interval. This resulted in a 35 kb region delimited by CER449813 and CER452353/54 containing the *DGR1* gene. Three residual recombinants were analyzed with two remaining InDel markers located in the 35kb area but no additional recombination events could be detected (figure 3.9).

3.3.1.3 IDENTIFICATION OF THE *DGR1* GENE

According to the gene annotations for chromosome II in the publicly available *Arabidopsis* genome database (www.arabidopsis.org; www.tigr.org), the 35 kb region comprising



FIGURE 3.4 GENOTYPING OF F_1 PLANTS RESULTING FROM RECIPROCAL CROSSES BETWEEN *dgr1* MUTANT PLANTS (COL-0) AND LER WT PLANTS.

Each pollination of a particular cross was checked by analyzing DNA from leaf material from one $F_{1(x)(y)}$ plant, resulting from the y^{th} pollination in a cross of a *Ler(x)* parent plant with a *dgr1(x)* parent, with the Col/Ler SLP marker nga8 (Bell and Ecker, 1994)¹. The concentrations of reagents used in the PCR reaction were: DNA template of a cTAB preparation (2µL/20µL PCR reaction volume); PCR buffer (50 mM potassium chloride, 10 mM Tris-HCl pH 9.0, 0.1% Triton X-100); 200 mM dNTPs (each); 1 mM nga8 forward primer (TGGCTTTCGTTTATAACATCC); 1µM nga8 reverse primer (GAGGGCAAATCTTATTCGG); 2µM MgCl₂; 0.5 U Taq polymerase for 20 ml PCR reaction volume. The following PCR profile was used: 1 minute at 94°C; 40 cycles: 30 seconds at 94°C; 30 seconds at 55°C; 30 seconds at 72°C; 4°C. PCR fragments were separated on a nusive agarose gel (3:1 solution in Tris Acetate EDTA (TAE) buffer) using ethidium bromide for visualization of the fragments. The obtained Ler and Col fragments have a length of 198bp and 154bp, respectively.

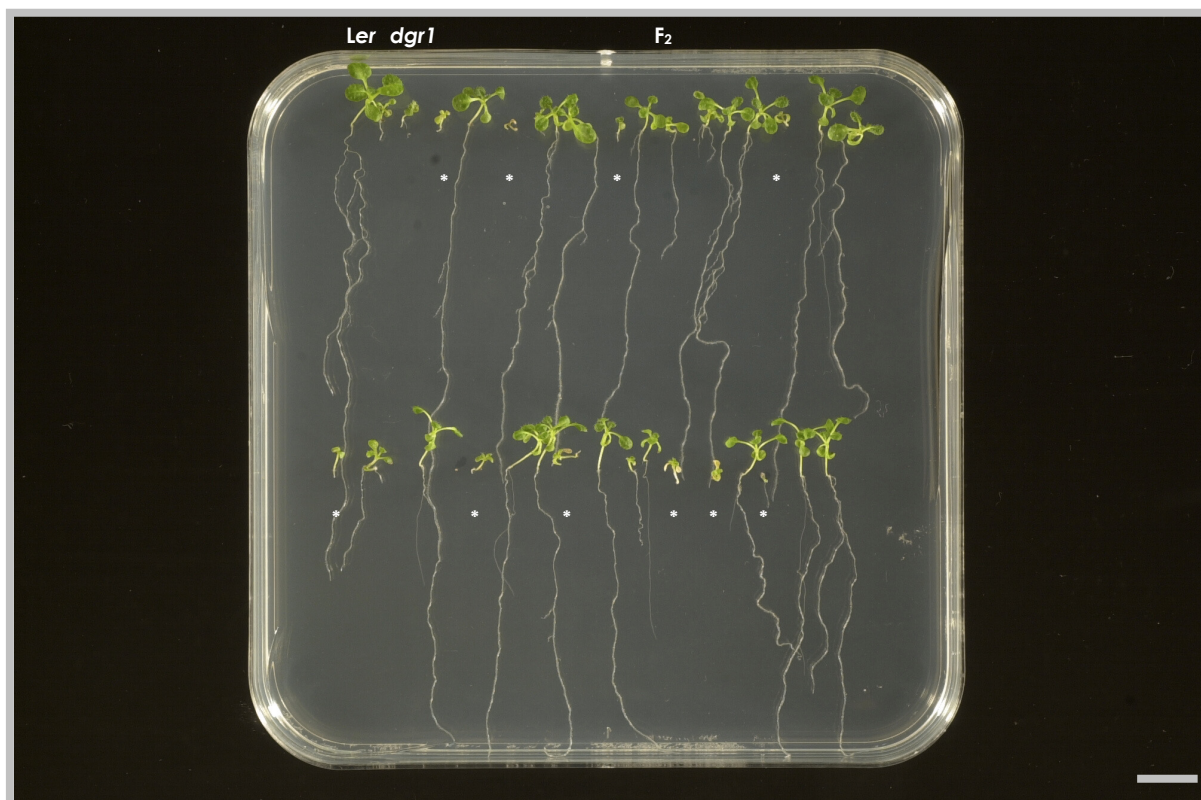


FIGURE 3.5 SEGREGATION OF AN F₂ MAPPING POPULATION DERIVED FROM A CROSS BETWEEN THE *Ler* WT AND THE *Col* *ts* MUTANT *dgr1* OF *ARABIDOPSIS THALIANA*. Seeds of the parental lines *Ler* and *dgr1*, and 32 F₂ seeds derived from the cross between the two parental lines were sown *in vitro* on standard MS growth medium. Plants were germinated at the permissive temperature of 18°C and then grown for 10 days at the restrictive temperature of 28°C. A 3:1 Mendelian segregation ratio of WT to mutant (*) plants is observed. Bar=1 cm.

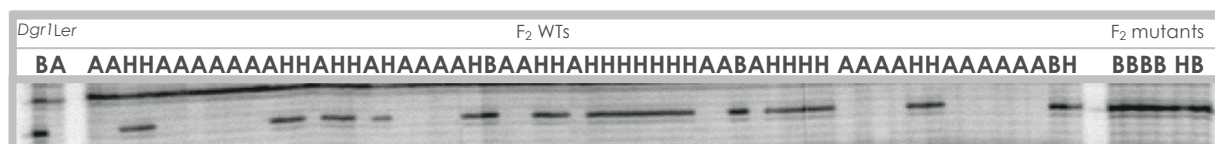


FIGURE 3.6 SEGREGATION PATTERN OF AN AFLP MARKER (SM240_448.2) SHOWING LINKAGE TO THE *dgr1* LOCUS. Since the *dgr1* mutation was generated in a *Col* background, complete linkage of an AFLP marker to the mutant locus is characterized by a segregation pattern showing a *Col* fragment (AFLP marker) that is in 25% of the F₂ individuals homozygous for absence, in 25% of the F₂ individuals homozygous for presence and in 50% of the F₂ individuals heterozygous. Genotypic data are reflected in intensities observed after visualization of the radio-actively labeled *Col* fragments using a PhosphorImager. These intensities are measured with the software AFLP-QuantarPro, which allows to assign a genotypic score to each intensity (A: homozygous for absence; B: homozygous for presence; H: heterozygous).

the *DGR1* gene encompasses 12 putative genes. To determine which of these 12 genes carried the *dgr1* mutation, all of them were sequenced in *dgr1* mutants and in *Col* WTs. By comparing the respective gene sequences from the mutants with those from the WTs, we detected two point mutations, designated *dgr1-1* (G₁₉A) and *dgr1-2* (G₁₀₁A), in the gene at2g45790 (*DGR1*) located in BAC

clone F4I18 (figure 3.9; figure 3.10). In the remaining 11 genes no mutations were detected. To further verify the presence of the two mutations, the at2g45790 gene was sequenced in 20 *Col* WT plants and in the 40 mutant recombinants, which were identified and used during the fine-mapping of *dgr1*. Both *dgr1-1* and *dgr1-2* were absent in the WTs and present in all 40 recombinants (data not shown).

TABLE 3.2 SET OF THIRTEEN ADDITIONAL PCs USED FOR THE AMPLIFICATION OF AFLP MARKERS LOCALIZED IN THE VICINITY AND WITHIN A 4.7 MB REGION ON THE LOWER HALF OF CHROMOSOME II CONTAINING THE *dgr1* LOCUS.

MARKER CODE ¹	SELECTIVE NUCLEOTIDES		AFLP MARKER(S) OF INTEREST
	SACI + 2	MSEI + 2	
SM26	AC	GC	SM26_495.4
SM33	AG	AA	SM33_202.4
SM34	AG	AC	SM34_231.6
			SM34_325.7
SM47	AG	TG	SM47_172.4/174.4 ¹
SM54	AT	CC	SM54_166.4
			SM54_318.8
			SM54_397.3
SM89	CC	GA	SM89_120.4
SM117	CT	CA	SM117_193.1/194.2 ¹
SM120	CT	CT	SM120_153.8
SM168	GG	CT	SM168_330.8
SM190	GT	TC	SM190_142.6
SM235	TG	GG	SM235_247.2
SM239	TG	TG	SM239_188.2
			SM239_289.7
SM244	TT	AT	SM244_283.8

¹Bi-allelic AFLP markers (figure 3.7).

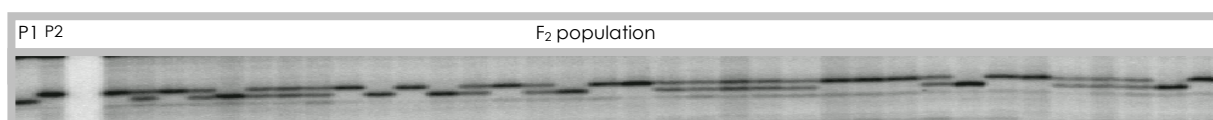


FIGURE 3.7 EXAMPLE OF THE F₂ SEGREGATION PATTERN OF A BI-ALLELIC AFLP MARKER. Bi-allelic AFLP markers are identified as a result of small insertions or deletions in the AFLP restriction fragments and occur at a low frequency. By its nature, a bi-allelic AFLP marker can be scored co-dominantly in a facile way by eye.

The *DGR1* gene consists of 2590 bp and is organized into 10 exons and 9 introns. The point mutations *dgr1-1* and *dgr1-2* both occur in the first exon (figure 3.9). Comparison of the WT and mutant cDNAs and their deduced amino acid sequences revealed that the missense mutation *dgr1-1* involves the substitution of a glycine residue (GGA) into an arginine residue (codon AGA), whereas the *dgr1-2* point mutation results in replacement of an arginine (codon CGA) by a glutamine (codon CAA) (figure 3.10). The deduced DGR1 protein consists of 246 amino acids with a predicted molecular mass of 27,732 Dalton and a predicted iso-electric point of 5.2.

3.3.1.4 *DGR1* ENCODES AN EUKARYOTIC PHOSPHOMANNOMUTASE

The *DGR1* gene (at2g45790) encodes an eukaryotic phosphomannomutase.

Phosphomannomutases are phosphotransferases belonging to the recently discovered haloacid dehalogenase (HAD) superfamily of enzymes (Koonin et al., 1994; Aravind et al., 1998). Although sequence identity across the entire HAD family is less than 15%, members share three conserved sequence motifs. They furthermore possess a conserved core

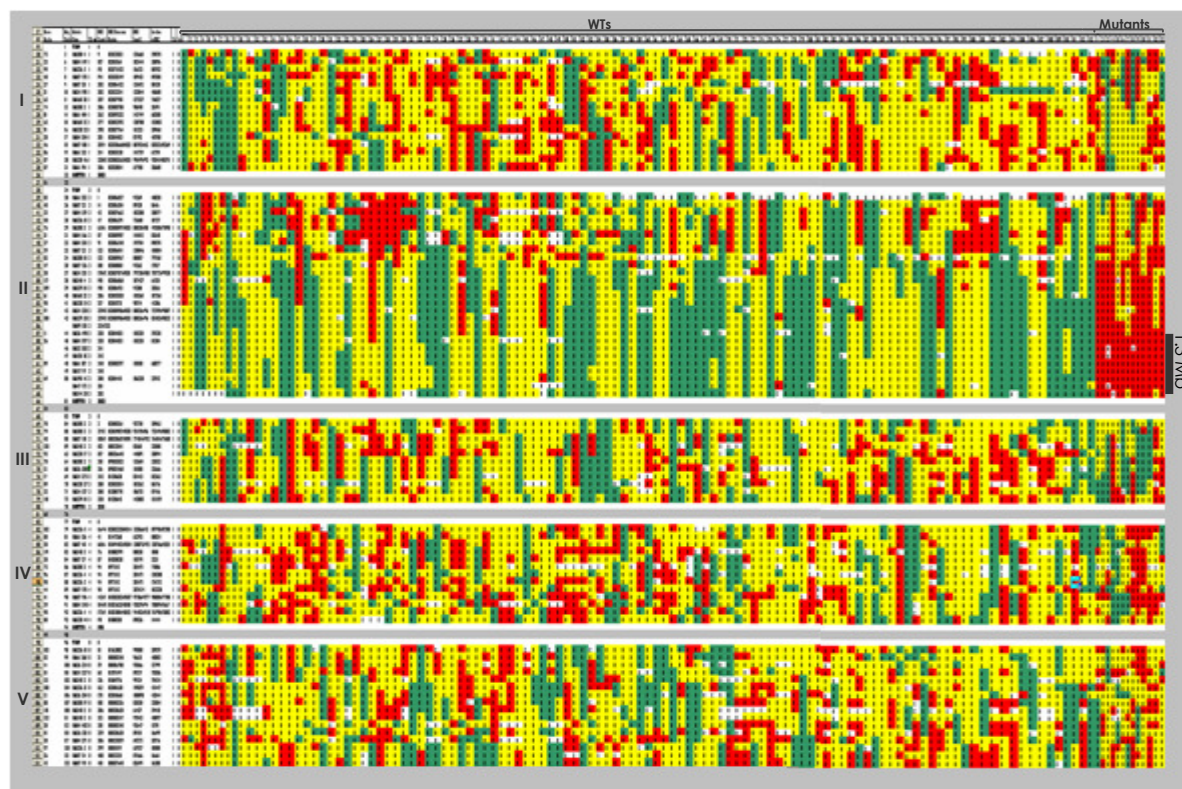


FIGURE 3.8 LINKAGE OF THE *dgr1* LOCUS WITH COL/LER AFLP MARKERS AT THE BOTTOM OF CHROMOSOME II AND NON-LINKAGE WITH COL/LER AFLP MARKERS ON CHROMOSOMES I, III, IV AND V. 126 F₂ plants, consisting of 113 WT and 13 mutants, together with the Ler and Col parents, were analyzed using in a first step a set of eight PCs (table 3.1). After scoring the resulting 85 AFLP markers, linkage to chromosome II and non-linkage to chromosomes I, III, IV and V were observed. In a next step, 13 additional PCs, amplifying AFLP markers in the vicinity and within the defined region of interest (lower half of chromosome II), were used for further linkage analysis performed on the same mapping population (table 3.2). This way, the region containing the *dgr1* locus could be narrowed down to 1.3 Mb (grey vertical bar) with SM26_495,4 as the upper flanking AFLP marker. Scoring of the markers (A: homozygous for absence; B: homozygous for presence; H: heterozygous) is performed using the software AFLP-Quantar Pro. Markers and their corresponding scores are then sorted according to their physical position on the different chromosomes. By assigning a different color to each score (A: —; B: —; H: —), linkage to a chromosome 2 genomic region could easily be detected.

domain and are grouped into three subfamilies (I, II, III) based on the presence and the location of a second domain, the 'cap' domain (Morais et al, 2000).

BLAST searches revealed the existence of several homologues for DGR1 in other plants species, such as *Oryza sativa* and *Medicago trunculata*. High sequence identity orthologues were also found in yeasts, such as *Schizosaccharomyces pombe*, *Candida albicans* and *Candida glabrata*, and animals, such as *Mus musculus*, *Rattus norvegicus*, *Macaca fasciculata*, *Drosophila melanogaster*, *Xenopus laevis* and many others. Furthermore, DGR1 shows 56% and 60% amino acid sequence

identity when compared to human phosphomannomutase I and human phosphomannomutase II, respectively (table 3.4; figure 3.11).

3.3.2 MAP-BASED CLONING OF *dgr2*

The *dgr2* mutant was isolated in a similar way as *dgr1*. The isolation and phenotypic characterization of the mutant have been described in detail in the previous chapter (chapter 2; table 2.1). F₂ mapping populations

TABLE 3.3 COL/LER INDEL POLYMORPHISMS USED FOR THE MAP-BASED CLONING OF THE *dgr1* LOCUS.

INDEL NAME ¹	BAC	FORWARD PRIMER (5'→3')	REVERSE PRIMER (5'→3')	FRAGMENT LENGTH (BP)		INDEL SIZE
				COL	LER	
CLONE NAME						
CER452389	F4L23	tggtgatgggccttttaac	gaacgtgtcactcggctttt	204	173	+31/-31
CER452385	F4L23	aacaaacgcctttccggtta	tccttgctcgagaagattaga	300	294	+6/-6
CER452369	F4L23	tcagtggtgttccccaatca	taaacgcttggcagttaggc	364	359	+5/-5
CER449803	F17K2	ttcgctctttcatggcagtt	gctagtgtgaacgcctcaaaa	214	206	+8/-8
CER449805	F17K2	tttgcgatgaactctggatca	tcctttaaaacaccactttcg	224	217	+7/-7
CER449809	F17K2	atatttacaattattgtttc	tgataatgattctagctaac	223	220	+3/-3
CER449810	F17K2	gaattgatatataggcaaga	tcataaagaataaacctgca	173	170	+3/-3
CER449813	F17K2	aacaacaaaacaaccagaaa	gaaagaaaagaccaaacatc	259	251	+8/-8
CER452357	F4I18	agatactgaggaggctgagg	ttgctgtggaacggagtc	220	214	+6/-6
CER452356	F4I18	gcaagagctttggactcgac	ggcgaagcaaaaacaagagtc	315	313	+2/-2
CER452353/54	F4I18	ccatgaaatgcccaatttfc	tgtggaaatgttagggaaaaga	229	225	+4/-4
CER452350	F4I18	tggatcgcctaatgaccaca	ccaccaaccaatattcactcaa	208	197	+11/-11
CER452347	F4I18	gccttcagcattcgggaacc	catgcatttaattacacatgacgtt	204	161	+43/-43
CER452346	F4I18	gggaataagcatttcggaga	tggatcattgtacgttaatttc	243	239	+4/-4
CER460503	T3F17	agataggcatgagcacagta	tcgcattcgaactccaaagtc	350	342	+8/-8
CER448608	F11C10	tgttgttagtaacgatggcaaa	ccacttcggctcttcgaac	420	414	+6/-6
CER449268	F14M4	ccacttcggtcttctcgaac	tttgtgaactattaaatcagatgttga	105	101	+4/-4
CER449257	F14M4	ttftagcgcagcatgaaagc	ccacatcaatccctctccat	458	442	+16/-16
CER449256	F14M4	cccgtacgtatttgtattatgc	ggcgagggatcagaaaatgt	136	116	+20/-20

¹Cereon Arabidopsis Polymorphism Collection

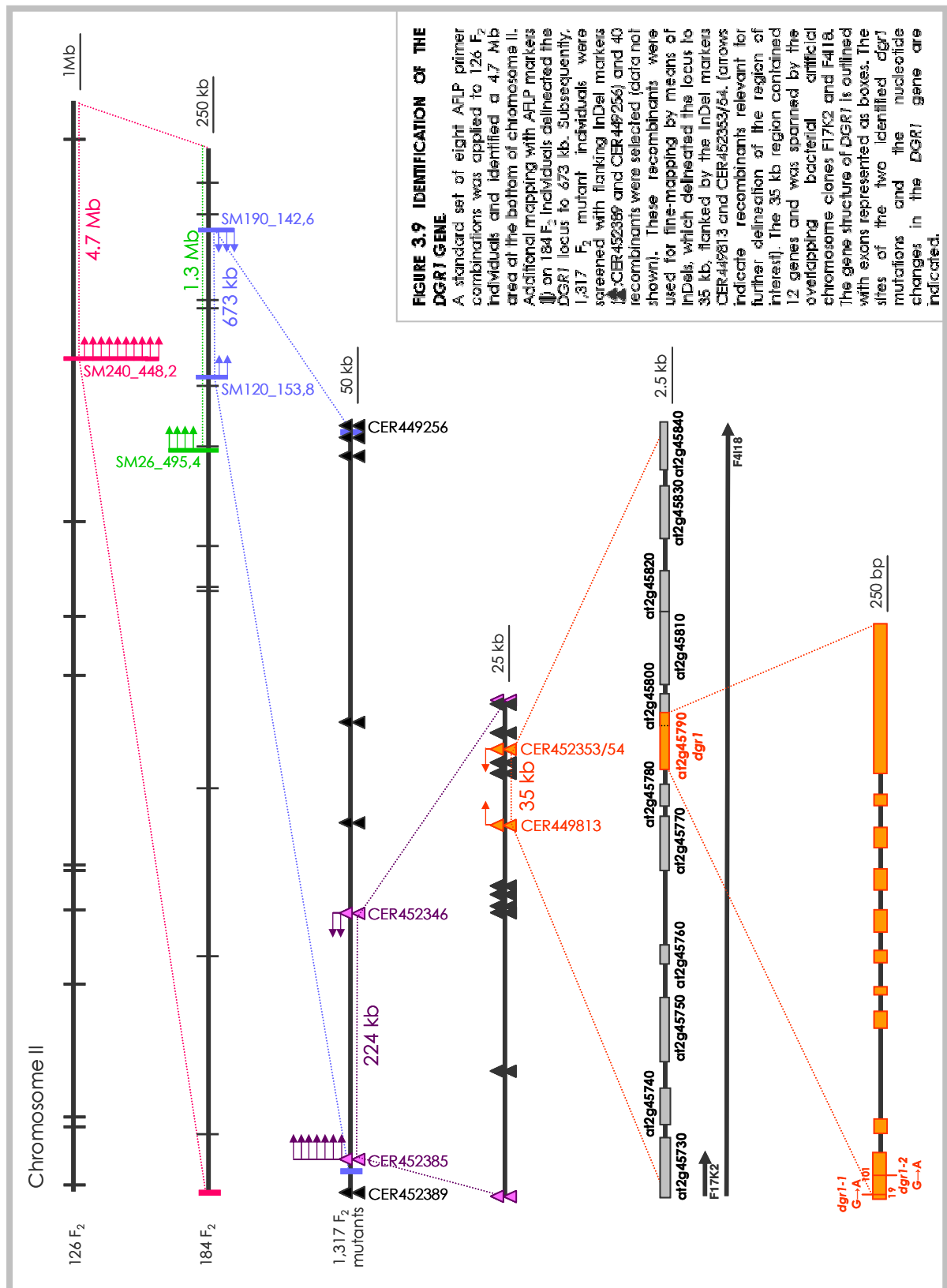
were obtained in a similar way as for *dgr1* (see paragraph 3.3.2; figure 3.12 and 3.13).

3.3.2.1 WHOLE-GENOME MAPPING OF *dgr2*

For the determination of the chromosomal segment containing the *dgr2* mutation, a total of 278 F₂ plants were analyzed, consisting of 211 WTs and 67 mutants. Segregation analysis of the 85 AFLP markers showed clear linkage to a genomic region on chromosome II (figure 3.12). The length of the chromosomal segment containing the gene of interest was 3.5Mb and delineated by the lower flanking marker SM61_322.0 and the upper end of chromosome II. The 3.5Mb interval consisted of a sequenced 38.4 kb interval, followed by a large non-sequenced area, comprising the nucleolar organizer region (NOR) and the upper telomere of chromosome II (figure 3.14).

3.3.2.2 *Dgr2* IS LOCALIZED IN THE NOR REGION OF CHROMOSOME II

The identified genomic interval comprises the NOR region and telomere of chromosome II and largely lacks sequence information due to its highly repetitive nature. This made further mapping of the *dgr2* locus unfeasible. At the position of the flanking marker SM61_322.0, 33 recombinants could still be identified, suggesting a genetic distance of 11.9cM or a physical distance of ~2.7Mb between this marker and the mutant locus. The 3.5Mb region contained only a very small sequenced 38.4 kb fragment preceding the NOR region. According to annotations in the *Arabidopsis* genome database, this fragment comprises two ribosomal rRNAs, six retrotransposons or pseudogenes and only one protein-coding gene, at2g01031 (figure 3.14). This gene encodes a hypothetical protein, functioning in the chloroplasts. To exclude the possibility that the ts *dgr2* mutation resides in at2g01031, we determined the sequence of this gene in the *dgr2* mutant and in the WT. However, based on the sequence comparison no mutation was detected.



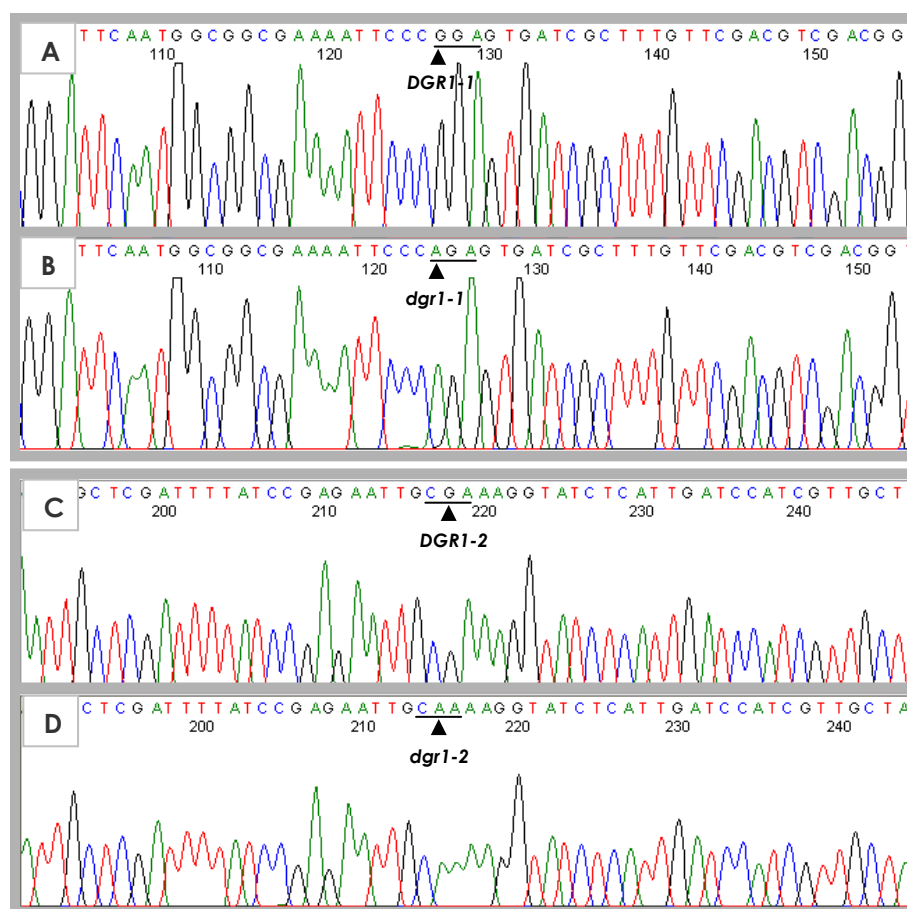


FIGURE 3.10 DNA SEQUENCE CHROMATOGRAMS OF *DGR1* MUTANT AND COL WT DNA SHOWING THE TWO POINT MUTATIONS, *dgr1-1* AND *DGR1-2* PRESENT IN THE MUTANT GENOME. After identification of a 35 kb region comprising the *DGR1* gene, the twelve genes localized within the interval were sequenced both in Col WTs (A, C) and in *dgr1* mutants (B, D). By comparing the gene sequences from the WTs with those from the mutants, we detected two point mutations, designated *dgr1-1* (G₁₉A transition) and *dgr1-2* (G₁₀₁A transition), in the gene at2g45790 (*DGR1*). The point mutations *dgr1-1* and *dgr1-2* both occur in the first exon. Comparison of the WT and mutant cDNAs and their deduced amino acid sequences revealed that the missense mutation *dgr1-1* involves the substitution of a glycine residue (codon GGA) into an arginine residue (codon AGA), whereas the *dgr1-2* point mutation results in replacement of an arginine (codon CGA) by a glutamine (codon CAA).

3.4 DISCUSSION

Using a genome-wide AFLP-based mapping approach (Peters et al., 2004), we mapped two ts mutations, *dgr1* and *dgr2*, affecting growth and development in *Arabidopsis thaliana*, to the lower and the upper arm of chromosome II, respectively. We succeeded in further fine-mapping of *dgr1* using individual InDel markers derived from the Cereon *Arabidopsis* Polymorphism collection (<http://www.arabidopsis.org/Cereon/index.html>) and found that the *DGR1* gene corresponds to at2g45790 (figure 3.9), a gene coding for eukaryotic phosphomannomutase in *Arabidopsis*.

3.4.1 THE TEMPERATURE-SENSITIVE PHENOTYPE OF *dgr1* IS DUE TO A DEFICIENCY IN PHOSPHOMANNOMUTASE ACTIVITY

Arabidopsis DGR1/at2g45790 codes for eukaryotic phosphomannomutase, a phosphotransferase belonging to the recently discovered superfamily of haloacid dehalogenases (HAD) (Koonin et al., 1994; Aravind et al., 1998). This family of enzymes comprises a large and diverse set of hydrolases found in the genomes of *Bacteria*, *Archaea* and *Eukarya*. Besides phosphotransferases, the HAD superfamily includes phosphatases, phosphonates, P-type ATPases, and dehalogenases, which are involved in cellular processes ranging from amino acid biosynthesis to detoxification (table 3.5).

Organism	Source	Length	Score	Identity%	Positives%	E-val
<i>A. thaliana</i>	Probable phosphomannomutase (EC 5.4.2.8) (PMM).	246	1289	100	100	1.5e-130
<i>O. sativa</i>	OSJNBa0032F06.16 protein.	248	1081	81	91	1.7e-108
<i>M. trunculata</i>	HAD-superfamily hydrolase subfamily IIB.	247	1062	80	91	1.7e-106
<i>S. pombe</i>	Phosphomannomutase (EC 5.4.2.8) (PMM).	257	819	62	77	9.7e-81
<i>X. laevis</i>	MGC85250 protein.	246	796	64	76	2.6e-78
<i>M. fasciculata</i>	Phosphomannomutase 2 (EC 5.4.2.8) (PMM 2).	246	781	60	77	1.0e-76
<i>H. sapiens</i>	Phosphomannomutase 2 (EC 5.4.2.8) (PMM 2).	246	780	60	77	1.3e-76
<i>N. crassa</i>	Probable phosphomannomutase.	273	780	59	73	1.3e-76
<i>D. discoideum</i>	Hypothetical protein.	249	780	60	75	1.3e-76
<i>G. zeae</i>	Hypothetical protein.	270	775	60	73	4.4e-76
<i>C. globosum</i>	Hypothetical protein.	265	771	56	73	1.2e-75
<i>G. domesticus</i>	Hypothetical protein.	247	771	61	73	1.2e-75
<i>C. albicans</i>	Phosphomannomutase (EC 5.4.2.8) (PMM).	252	766	60	76	4.0e-75
<i>C. albicans</i>	Hypothetical protein PMM1.	252	766	60	76	4.0e-75
<i>D. hansenii</i>	Similar to sp P31353 Candida albicans CaPMM1 phosphomannomutase.	252	766	61	76	4.0e-75
<i>M. musculus</i>	Phosphomannomutase 2 (EC 5.4.2.8) (PMM 2).	242	760	60	77	1.7e-74
<i>M. musculus</i>	Adult male cerebellum cDNA, RIKEN full-length enriched library, clone:1500000C17 product:phosphomannomutase 2, full insert sequence (CDNA, RIKEN full-length enriched library, clone:10C0037F16 product:phosphomannomutase 2, full insert sequence) (2 days neonate sympathetic ganglion cDNA, RIKEN full-length enriched library, clone:7120484K22 product:phosphomannomutase 2, full insert sequence) (Phosphomannomutase 2) (Adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430404G24 product:phosphomannomutase 2, full insert sequence).	242	760	60	77	1.7e-74
<i>M. musculus</i>	18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110002P07 product:phosphomannomutase 2, full insert sequence.	242	759	60	77	2.2e-74
<i>C. glabrata</i>	Similar to sp P07283 Saccharomyces cerevisiae YFL045c SEC53 phosphomannomutase.	255	751	60	73	1.6e-73
<i>C. neoformans</i>	Hypothetical protein.	271	749	56	76	2.5e-73
<i>C. neoformans</i>	Phosphomannomutase, putative.	271	749	56	76	2.5e-73
<i>K. lactis</i>	Phosphomannomutase (Similarity).	254	749	59	73	2.5e-73
<i>Y. lipolytica</i>	Similar to sp P31353 Candida albicans Phosphomannomutase.	253	747	59	73	4.1e-73
<i>L. mexicana</i>	Phosphomannomutase (EC 5.4.2.8).	247	745	59	74	6.7e-73
<i>L. major</i>	Phosphomannomutase, putative (EC 3.1.3.11).	247	741	59	74	1.8e-72
<i>B. taurus</i>	Phosphomannomutase 2.	246	741	58	75	1.8e-72
<i>A. gossypii</i>	ABR236Wp.	253	740	57	74	2.3e-72
<i>B. rerio</i>	Hypothetical protein pmm2.	250	738	58	75	3.7e-72
<i>X. laevis</i>	MGC82669 protein.	260	735	58	74	7.7e-72
<i>D. melanogaster</i>	Probable phosphomannomutase (EC 5.4.2.8) (PMM).	254	733	56	73	1.3e-71
<i>A. gambiae</i>	ENSANGP00000017981 (Fragment).	254	727	58	73	5.4e-71
<i>T. brucei</i>	Phosphomannomutase, putative (EC 5.4.2.8).	245	722	58	72	1.8e-70
<i>S. cerevisiae</i>	Phosphomannomutase (EC 5.4.2.8) (PMM).	254	720	58	70	3.0e-70
<i>A. nidulans</i>	Hypothetical protein.	473	719	55	71	3.8e-70
<i>X. laevis</i>	LOC495465 protein.	260	716	57	73	7.9e-70
<i>M. musculus</i>	Adult male kidney cDNA, RIKEN full-length enriched library, clone:0610040B01 product:phosphomannomutase 2, full insert sequence.	232	712	58	76	2.1e-69
<i>R. norvegicus</i>	Phosphomannomutase 1.	262	711	55	71	2.7e-69
<i>M. musculus</i>	Phosphomannomutase 1 (EC 5.4.2.8) (PMM 1).	262	708	57	73	5.6e-69
<i>D. pseudoobscura</i>	GA10496-PA (Fragment).	254	708	56	73	5.6e-69
<i>M. musculus</i>	Adult male lung cDNA, RIKEN full-length enriched library, clone:1200007E18 product:phosphomannomutase 1, full insert sequence (Blastocyst blastocyst cDNA, RIKEN full-length enriched library, clone:11C0027C23 product:phosphomannomutase 1, full insert sequence).	262	708	57	73	5.6e-69
<i>M. fasciculata</i>	Brain cDNA, clone: QtrA-10834, similar to human phosphomannomutase 1 (PMM1).	262	701	56	72	3.1e-68
<i>T. cruzi</i>	Phosphomannomutase, putative (EC 3.1.3.11).	245	698	54	71	6.4e-68
<i>U. maydis</i>	Hypothetical protein.	271	697	55	68	8.2e-68
<i>H. sapiens</i>	Phosphomannomutase 1 (EC 5.4.2.8) (PMM 1) (PMMH-22).	262	696	56	71	1.0e-67
<i>A. oryzae</i>	Phosphomannomutase.	269	694	54	69	1.7e-67
<i>C. parvum</i>	Phosphomannomutase [EC:5.4.2.8] (Fragment).	252	692	55	70	2.8e-67
<i>D. discoideum</i>	Similar to Candida albicans (Yeast). Phosphomannomutase (EC 5.4.2.8) (PMM) (Hypothetical protein).	222	684	60	77	2.0e-66
<i>C. neoformans</i>	Sec53p.	306	542	54	75	1.0e-62
<i>E. histolyticae</i>	Phosphomannomutase, putative.	243	648	52	70	1.3e-62
<i>A. locustae</i>	Phosphomannomutase.	249	633	51	69	5.0e-61

TABLE 3.4 OUTPUT OF A BLAST SEARCH ([WWW.EBI.AC.UK.HTML](http://www.ebi.ac.uk/html)) FOR THE ARABIDOPSIS DGR1 PROTEIN/EUKARYOTIC PHOSPHOMANNOMUTASE. An NCBI-BLAST2 Protein Database Query (software WU-BLAST version 2.0; EMBL-EBI) (50 best hits are shown) revealed the existence of several homologues for DGR1 in other species.

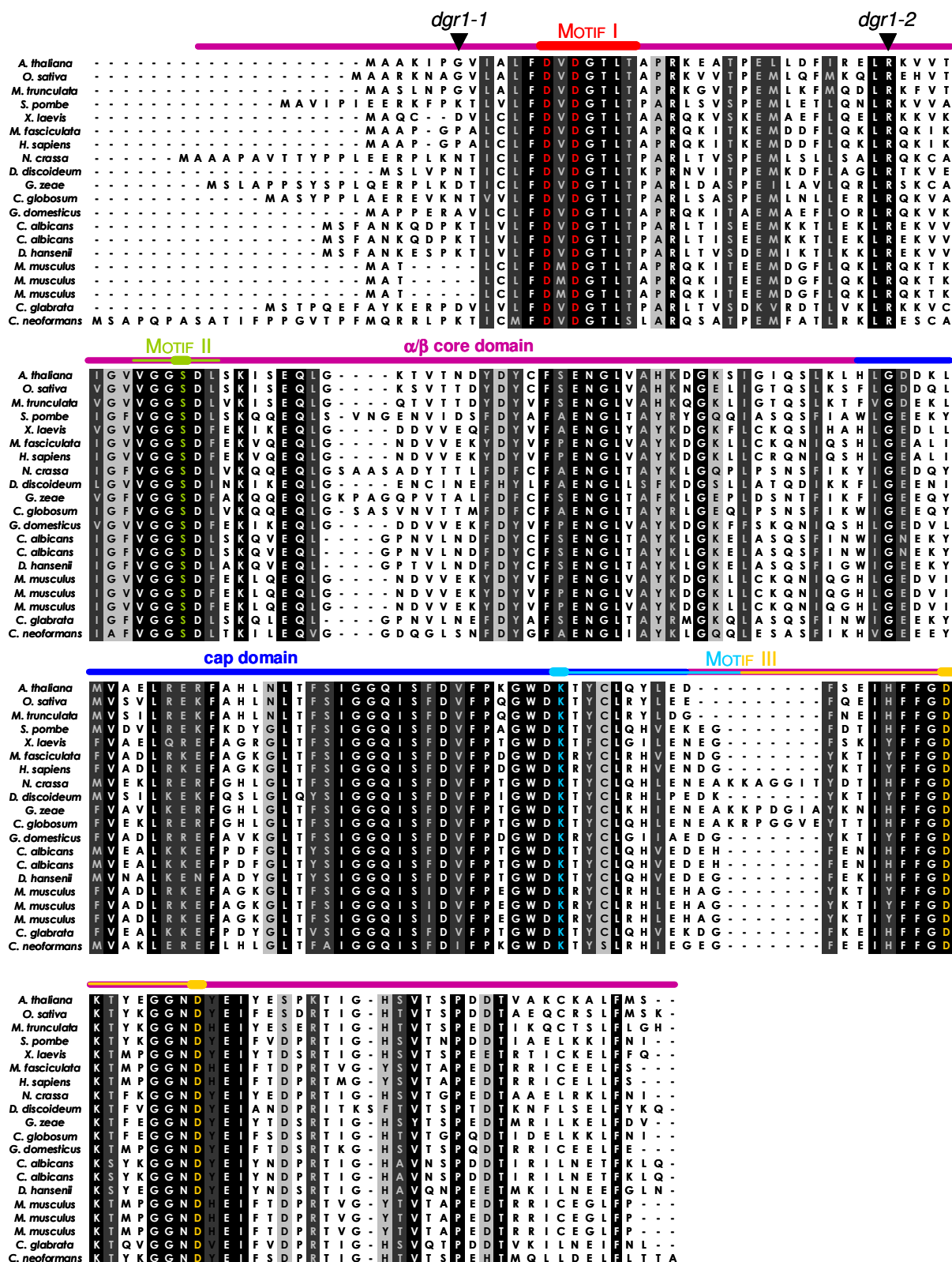


FIGURE 3.11 ALIGNMENT OF THE DEDUCED AMINO ACID SEQUENCE OF THE *ARABIDOPSIS DGR1* GENE AND SEVERAL OF ITS ORTHOLOGUES IN OTHER PLANT SPECIES, YEAST AND ANIMALS. Identical, conserved and semi-conserved residues are presented in black, in dark grey and in light grey, respectively. The α/β-core domain (—), the cap domain (—) and the three conserved motifs are indicated by lines above the sequence. Amino acids participating in the catalytic activity of the enzyme are coloured corresponding their positions in catalytic loop 1 (—), 2 (—), 3 (—) and 4 (—) of the core domain. Arrowheads indicate the positions of the mutations *dgr1-1* and *dgr1-2*.

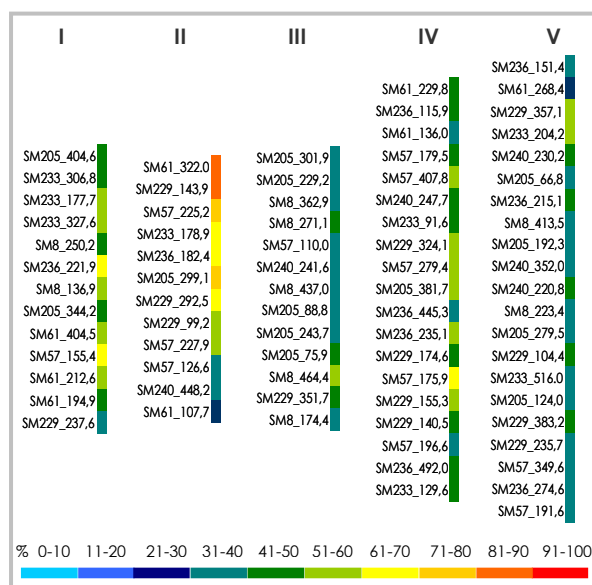


FIGURE 3.12 LINKAGE OF THE *DGR2* LOCUS WITH *Col/Ler* AFLP MARKERS AT THE TOP OF CHROMOSOME II AND NON-LINKAGE WITH *Col/Ler* AFLP MARKERS ON CHROMOSOMES I, III, IV AND V. 278 *F*₂ plants, consisting of 211 WTs and 67 mutants, together with their *Ler* and *Col* parents, were analyzed using eight PCs (table 3.1). The figure visualizes the abundance (in %) of each *Col* AFLP fragment in the population of 67 *F*₂ mutants.

We detected two point mutations in the *DGR1* (at2g45790) gene, designated *dgr1-1* (G₁₉A transition) and *dgr1-2* (G₁₀₁A transition) (figure 3.10). Comparison of the WT and mutant deduced amino acid sequences revealed that *dgr1-1* is a missense mutation involving the substitution of a glycine residue (codon GGA) into an arginine residue (codon AGA), whereas the *dgr1-2* point mutation results in the replacement of an arginine (codon CGA) by a glutamine (codon CAA). Since ts mutations typically cause a conformational change of the protein in the mutant at restrictive (usually higher) temperatures, it is plausible that *dgr1-1* and/or *dgr1-2* could affect the structure of the DGR1 protein.

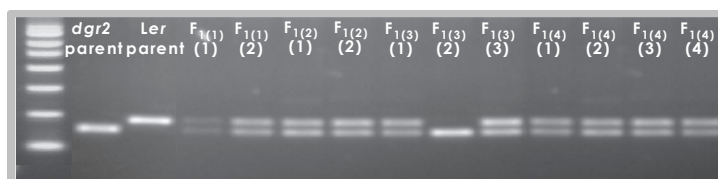
3.4.1.1 *Dgr1-1* AND *dgr1-2* CAUSE AMINO ACID SUBSTITUTIONS IN THE PERIPHERY OF THE PHOSPHOMANNOSE α/β -CORE DOMAIN

Phosphomannomutase is a homodimeric member of the magnesium-dependent phosphotransferase family of

HAD enzymes, which catalyze cellular phosphoryl group transfer reactions (Knowles et al., 1980; Mildvan, 1997; Frey et al., 1987; Dzeja et al., 2003). The thermodynamics of these reactions vary greatly (Voet and Voet, 1995) but the kinetic barrier is consistently high. Phosphotransferases enhance the reaction rate by lowering the energy barrier for the reaction via strong binding to a transition state (Lad et al., 2003).

Solution phosphoryl-transfer reactions can proceed by three, mechanistically distinct, nucleophilic substitution pathways (dissociative, concerted and associative) (figure 3.15). These differ in the timing of the formation of the phosphorus-nucleophile bond and cleavage of the phosphorus leaving-group bond. The prevailing chemical pathway is determined by the nature of the nucleophile, electrophile and leaving group, and by the solvent. There is evidence to support the view that all three types of chemical pathways are catalyzed by phosphotransferases, but the rules governing selection of a pathway for a given enzyme-reactant pair are not known (Catrina and Hengge, 2003; Grzyska et al., 2003; Anderson, 2001; Lahiri et al., 2003).

All members of phosphotransferase family have structures consisting of a typical α/β -core domain and a second cap domain, with the active site located at the domain-domain interface. The core domain, responsible for catalysis, has a highly conserved tertiary structure, while the cap domain is characterized by divergent sequences and sizes that confer specificity on the enzymes and result in the diverse substrate range (Wang et al., 2005; Morais et al., 2000). Furthermore, the catalytic mechanism of dephosphorylation in this family is believed to involve the formation of a phosphoaspartate-enzyme intermediate, which could be supported by the three conserved motifs in the active-site scaffold (figure 3.11; figure 3.16). Motif I has the sequence DXDX (T/V),



¹The concentrations of reagents used in the PCR reaction were: DNA template of a cTAB preparation (2μL/20μL PCR reaction volume); PCR buffer (50 mM potassium chloride, 10 mM Tris-HCl pH 9.0, 0.1% Triton X-100); 200 mM dNTPs (each); 1 mM nga8 forward primer (TGGCTTCGTTTATAACATCC); 1μM nga8 reverse primer (GAGGGCAAATCTTATTTCGG); 2μM MgCl₂; 0.5 U Taq polymerase for 20 ml PCR reaction volume. The following PCR profile was used: 1 minute at 94°C; 40 cycles: 30 seconds at 94°C; 30 seconds at 55°C; 30 seconds at 72°C; 4°C. PCR fragments were separated on a nusive agarose gel (3:1 solution in Tris Acetate EDTA (TAE) buffer) using ethidium bromide for visualization of the fragments. The obtained Ler and Col fragments have a length of 198bp and 154bp, respectively.

FIGURE 3.13 GENOTYPING OF F₁ PLANTS RESULTING FROM RECIPROCAL CROSSES BETWEEN DGR2 MUTANT PLANTS (COL-0) AND LER WT PLANTS. Each pollination of a particular cross was checked by analyzing DNA from leaf material from one F_{1(x)(y)} plant, resulting from the yth pollination in a cross of a Ler(x) parent plant with a dgr1(x) parent, with the Col/Ler SSLP marker nga8 (Bell and Ecker, 1994)¹. F₂ seeds were subsequently harvested from those F₁ plants that were confirmed to be heterozygous by the microsatellite analysis. The resulting segregating F₂ populations were used for the map-based cloning process (figure 3.5).

in which the first absolutely conserved aspartate residue forms an intermediate with the substrate (Collet et al., 1998). Motif II contains a conserved serine or threonine residue that forms a hydrogen bond with the phosphoryl oxygen atom (Wang et al., 2001). Motif III consists of a conserved lysine residue followed by a number of less conserved residues and a strictly conserved aspartate residue, which helps to coordinate the magnesium ion required for the enzyme activity (Morais et al., 2000; Wang et al., 2001). Examination of the sequence lengths of the cap domains shows that there are three distinct structural subgroups within the HAD family. The first subgroup includes proteins that contain an inserted cap domain between motif I and motif II, the second, to which phosphomannomutase belongs, includes proteins with a large inserted cap domain between motif II and motif III, and the third contains proteins with little to no inserted domain.

As depicted in figures 3.12 and 3.16, *dgr1-1* and *dgr1-2* result in amino acid substitutions in the periphery of the α/β-core domain. The changed amino acids will therefore not directly and specifically affect the regions that are essential for substrate binding (substrate-specific loop in cap domain) or catalytic activity (catalytic scaffold in core domain).

What might be expected when dealing with ts mutations is indeed a more subtle interference, since the effect is only observed under certain conditions.

Therefore, we suggest that, at elevated temperatures, the induced amino acid substitutions might disrupt the highly conserved tertiary structure of the core domain, which is responsible for catalysis. This is primarily supported by the fact that both mutations cause relatively drastic amino acid substitutions: the *dgr1-1* point mutation causes a change of a neutral aliphatic glycine residue into a positively charged hydrophilic arginine, whereas *dgr1-2* results in the substitution of a positively charged basic arginine into a negatively charged acidic glutamine. These changes of charge and/or hydrophilic nature of amino acid residues could affect interactions, needed to maintain the active enzyme conformation but also unbalance enzyme-solvent bonds, needed to maintain optimal enzyme activity. The first is often the case with buried residues, while the latter usually concerns exposed amino acids, positioned at the outer surface of the protein. Although both mutations could equally contribute to the ts phenotype, our results suggest that *dgr1-2* is possibly solely responsible. Support can be found in figure 3.11, which depicts the alignment of *Arabidopsis* phosphomannomutase to a number of its orthologues in other plant species, yeast and animals. It is clear that glycine at position 7 (*dgr1-1*) is highly variable, whereas arginine at position 37 (*dgr1-2*) is conserved in all shown species. More extensive BLAST searches revealed that in rare cases the arginine residue is replaced by a lysine, a similar positively charged basic residue.

FIGURE 3.14 THE DGR2 GENE WAS MAPPED TO A CHROMOSOMAL REGION OF 3.5 MB ON THE TOP OF CHROMOSOME II, CONSISTING OF A SEQUENCED (RED) 38.5 KB REGION, THE NOR2 (BLUE) AND THE UPPER TELOMERE (DARK BLUE).

(blue) and the upper telomere (dark blue). Grey lines represent AFLP markers used for the mapping of *dag2* on the chromosomal segment. The diagram at the bottom of the figure depicts the tandem arrangement of rRNA genes in the NOR region and the organization of one gene in the array.

TABLE 3.5 PROMINENT MEMBERS OF THE HALOACID DEHALOGENASE FAMILY SORTED BY CLASS AND REACTION TYPE

Enzyme	Class	Substrate
C-Cl cleavage		
2-haloalkanoic acid dehalogenase	I	2-haloalkanoic acid
C-P cleavage		
Phosphonoacetaldehyde hydrolase	I	Phosphonoacetaldehyde
CO-P cleavage		
Phosphoserine phosphatase	I	L-Phosphoserine
Mitochondrial 5'-(3')deoxyribonucleotidase	I	dUTP
Sucrose-6 ^F -phosphate phosphatase	II	Sucrose-6 ^F -phosphate
Mg ²⁺ -dependent phosphatase (MDP1)	III	Protein phosphotyrosine
8KDO phosphatase	III	3-Deoxy- α -manno-octulosonate 8-phosphate
CO-P cleavage/formation (mutase)		
β -phosphoglucosmutase	I	β -Glucose 1-phosphate
Phosphomannomutase	II	α -Mannose 1-phosphate
PO-P cleavage		
Sarcoplasmic Ca ²⁺ -ATPase	I	ATP
Cu ²⁺ /H ⁺ -ATPase	I	ATP

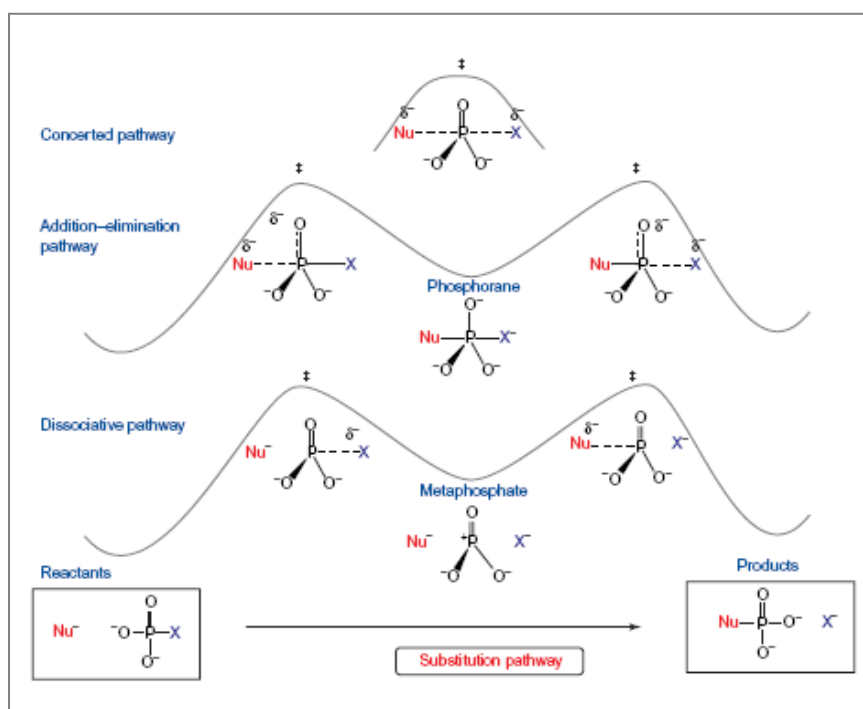


FIGURE 3.15 THREE POSSIBLE CHEMICAL MECHANISMS FOR PHOSPHORYL TRANSFER BETWEEN TWO NUCLEOPHILES (Nu) DIFFER IN THE TIMING OF BOND FORMATION AND CLEAVAGE (THE SUBSTITUTION PATHWAY) (ALLEN AND DUNAWAY-MARIANO, 2004). A hypothetical free energy diagram for each path is shown in the background. These highlight the fact that in the case of the addition-elimination (associative) and dissociative mechanisms an intermediate is formed, whereas there is no intermediate in the concerted reaction. In the addition-elimination mechanism, bond making to form a pentavalent phosphorane intermediate precedes bond breaking. In the dissociative mechanism, bond breaking to form a metaphosphate intermediate precedes bond making. In the concerted reaction, bonds are simultaneously made and broken in a single step, and the transition state resembles either metaphosphate or phosphorane.

Using the software tool SIFT ('Sorting Intolerant From Tolerant'), which predicts, based on sequence homology, whether amino acid substitutions will be tolerated or deleterious for protein function, our empiric observations were confirmed (Pauline and Henikoff, 2001; 2003; <http://blocks.fhcrc.org/sift/SIFT.html>).

The output also indicated the existence of natural DGR1 homologues, carrying an arginine at position 7.

Based on these arguments that are supported by the results, we suggest that the ts phenotype of *dgr1* is due to a conformational change in the tertiary structure of the phosphomannomutase α/β -core domain at elevated temperatures. Furthermore, we strongly

FIGURE 3.16 THE CATALYTIC SCAFFOLD IN THE HALOACID DEHALOGENASE (HAD) FAMILY OF PHOSPHOTRANSFERASES (ALLEN AND DUNAWAY-MARIANO, 2004). (A) Schematic of the roles of the four loops of the core domain that comprise the catalytic scaffold. The activity 'switch' is located on loop 4 of the catalytic scaffold (yellow), which positions one carboxylate residue to function as a general base for the dehalogenases and either two or three carboxylates to bind the Mg^{2+} cofactor, essential for the phosphotransferases. CO represents the backbone carbonyl oxygen of the moiety that is two residues downstream from the loop 1 nucleophile (red). The side-chain at this position is also used as an acid-base catalyst by phosphatase and phosphomutase HAD members. Loop 2 (green) and 3 (cyan) serve to position the nucleophile and substrate phosphoryl moiety. (B) Ribbon diagram (catalytic residues core (magenta); domain: loop 1, red; loop 2, cyan; loop 3, green; loop 4, yellow; cap domain (blue): substrate specificity loop (SSL), soft blue) of a predicted 3D model for the monomeric subunit of *Arabidopsis* phospho-mannomutase. Amino acid substitutions caused by *dgr1-1* (1) and *dgr1-2* (2) are highlighted (<http://www.ncbi.nlm.nih.gov/Structure>)



believe that this conformational change is possibly solely caused by the substitution of the positively charged basic arginine residue at position 37, which is proven to be highly conserved among species, by a negatively charged acidic glutamine residue.

3.4.1.2 BIOLOGICAL FUNCTIONS OF THE *DGR1* GENE ENCODING PHOSPHOMANMUTASE AND EFFECTS OF DELETERIOUS MUTATIONS

Phosphomannomutase catalyzes the interconversion of D-mannose-6-phosphate and D-mannose-1-phosphate

and is an enzyme that participates in cell wall biosynthesis and protein glycosylation. Furthermore, it is believed to catalyze a step in the plant ascorbic acid biosynthetic pathway proposed by Wheeler et al. (1998) (figure 3.17). Despite the importance of these processes, the plant enzyme has only been partially purified in a few species, such as *Amorphophallus konjac* and *Zea mays*.

According to the *Arabidopsis* microarray database Genevestigator (<https://www.genevestigator.ethz.ch/at/>), at2g45790 shows a moderate expression in the entire plant. Only in the primary root, the gene is highly expressed. The *DGR1* gene is highly conserved among plants (~90% amino acid sequence similarity) and between plants and other eukaryotes (~

75% amino acid sequence similarity) (table 3.4; figure 3.11). It is a single copy gene, as can be expected from a gene with an essential function, and hence has no paralogues. Therefore, we can reasonably assume that the *DGR1* gene encodes a phosphomannomutase protein, which enzymatically converts D-mannose-6-phosphate to D-mannose-1-phosphate. The essential nature of the *DGR1* gene is consistent with the essential role of the activated form of mannose-1-phosphate, GDP-mannose, a sugar that functions in a number of diverse cellular processes.

Cell wall biosynthesis

Both prokaryotes and eukaryotes utilize GDP-mannose in the synthesis of complex structural carbohydrates. Mannose is a component of microbial lipopolysaccharides and exopolysaccharides. Similarly, GDP-mannose also contributes to the synthesis of three different structural carbohydrates in plant cell walls. First, hemicellulose polymers such as glucomannans and galactomannans, contain D-mannose that is derived from its activated form. Mannans are major wall-based carbohydrate stores in some seeds (Bewley et al., 1997). Second, GDP-mannose is the substrate for GDP-D-mannose-4,6-dehydratase, an enzyme that catalyzes the first step in GDP-L-fucose biosynthesis, which is encoded by the *MUR1* gene of *A. thaliana* (Bonin et al., 1997). L-Fucose is present in both plant cell walls and on glycoproteins. Finally, L-galactose, a minor component of plant cell walls, is synthesized from GDP-mannose (Wheeler et al., 1998).

Protein Glycosylation

In addition to a major function in structural carbohydrate biosynthesis, mannose also plays a key role in protein glycosylation in eukaryotes. D-mannose, the major carbohydrate component of

both N- and O-linked glycoproteins, is derived from GDP-mannose during the glycosylation process. Glycoprotein mutants with known or predicted phosphomannomutase deficiencies have been described in yeast and humans.

In yeast, a temperature-sensitive mutant, generated by gene disruption of the *SEC53* gene encoding phosphomannomutase, was arrested in growth at the restrictive temperature (Kepes and Schekman, 1988).

In humans, a recessive genetic disease referred to as carbohydrate-deficient glycoprotein syndrome type Ia (CDG-Ia; Jaeken syndrome) results in underglycosylation of secretory proteins, lysosomal enzymes and membrane glycoproteins due to a defect in lipid-linked oligosaccharide biosynthesis (Powell et al., 1994; Krasnewich et al., 1995). It has been reported that cells from patients with this disease are severely deficient in phosphomannomutase activity (Van Schaffingen and Jaeken, 1995). Phosphomannomutase has two isozymes in humans with 66% sequence homology: PMM1 and PMM2. They catalyze the same reaction at approximately equal rates, but PMM2 is more substrate-specific (Pirard et al., 1997; 1999). PMM2 is the defective enzyme in CDG-Ia. The clinical features vary with age, and usually go through four stages: the infantile alarming multisystem stage, the childhood ataxia-mental retardation stage, the teenage leg atrophy stage and the adult hypogonadic stage. The overall mortality rate is approximately 20%, mostly occurring in the first years of life due to infection, multiple organ failure or heart tamponade (Jaeken et al., 1991; 2001). Up to now, more than seventy mutations in the *PMM2* gene, resulting in CDG-Ia have been identified, most of them being of the missense type. The absolute absence of patients harbouring two null alleles points to an essential role of PMM2 in normal cellular functions, and suggests that such zygotes are lost in early pregnancy.

L-Ascorbic Acid Biosynthesis

Phosphomannomutase also plays a key role in plant L-ascorbate biosynthesis by converting D-mannose-6-phosphate to D-mannose-1-phosphate in the Smirnoff/Wheeler pathway (figure 3.17 step 3). In plants, a diversity of physiological roles can be ascribed to the antioxidant and cellular reductant vitamin C (L-ascorbic acid) (Smirnoff, 1996).

We suggest that a possible ascorbate deficiency in *dgr1* causes accumulation of reactive oxygen species at the higher temperature, leading to a stress-sensitive phenotype.

Decreased growth, observed in the *dgr1* mutant, was also seen in the *vtc-1* mutant, an ozone-sensitive ascorbic acid mutant, deficient in GDP-mannose pyrophosphorylase (Conklin et al., 1996; 1997; figure 3.17 step 4), and might reflect an effect on the supply of C-skeletons for cell wall biosynthesis. On the other hand, in view of the many correlations between ascorbate and growth and the putative role of ascorbate in the plant cell cycle, part of the effect of low ascorbate could also be due to the disruption of control mechanisms involved in cell division and/or elongation (Noctor and Foyer; 1998a). One mechanism may be accelerated cross linking of cell wall components, which limits cell expansion in the mutant. This notion is supported by observations of Veljovic-Jovanovic et al. (2001) that show a substantial increase in total leaf peroxidase. In addition to the effect of peroxidase induction, the lower extracellular ascorbate concentration in the mutant creates an environment that markedly favors cross linking.

Given the many varied roles of ascorbic acid both as an anti-oxidant and a cofactor, one presumes that mutant plants totally devoid of ascorbic acid would be inviable. This is what we observe for the *dgr1* mutant at the restrictive temperature and this was also seen with the embryo-lethal *cyt1-2* mutant (allelic with *vtc-1*), which contained an

embryonic lethal mutation that creates a stop codon in the GDP-D-mannose pyrophosphorylase gene (Nickle and Meinke, 1998). However, as mutations such as *dgr1* and *cyt1-2* are predicted to be pleiotropic, it is difficult to say whether cell wall defects, protein glycosylation alterations and/or ascorbic acid deficiency causes the lethality.

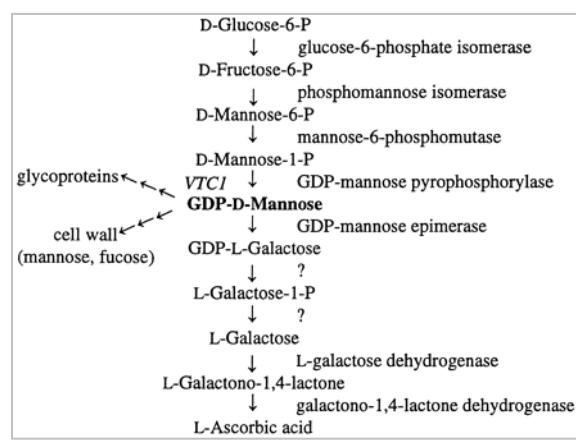


FIGURE 3.17 PROPOSED PATHWAY FOR ASCORBIC ACID BIOSYNTHESIS IN HIGHER PLANTS (WHEELER ET AL., 1998). The enzymatic step catalyzed by the *VTC1* gene product is shown as well as branchpoints to the synthesis of cell wall carbohydrates and glycoproteins. Enzymes: GPI, phosphoglucose isomerase; PMI, phosphomannose isomerase; PMM, phosphomannomutase; GMPP, GDP-D-mannose pyrophosphorylase; GME, GDP-D-mannose-3,5-epimerase; GDH, L-galactose dehydrogenase; GLDH, L-galactono-1,4-lactone dehydrogenase.

3.4.2 THE *dgr2* MUTATION IS LOCALIZED IN THE UPPER SUBTELOMERIC REGION OF CHROMOSOME II

Using a genome-wide mapping approach, a 3.5Mb region containing the *DGR2* gene was defined at the top of chromosome II. Due to the lack of informative polymorphisms in this region that discriminate ecotype-specific alleles in a mapping population, the mapping of the *DGR2* gene could not be continued.

Basically, the chromosomal segment at the top of chromosome II, comprising *dgr2*, consists of four distinct regions: a very small sequenced region

(38.5kb), the non-sequenced NOR2 (~3.5Mb), the NOR2-telomere junction and the upper telomere (~2-5kb).

First of all, it is highly improbable that the mutant locus resides within the small 38.5kb sequenced fragment. For the flanking marker SM61_322.0, which delineates the 3.5Mb region, 33 recombinants (within a population of 278 F₂ individuals) were identified, implying a genetic distance of 11.9cM or a physical distance of 2.7Mb between this marker and the mutant locus. Furthermore, the sole protein-coding, potential candidate gene in this region, at2g01031, was sequenced and proved to be free of mutations.

A more obvious reasoning would be that *dgr2* is localized in NOR2, which largely occupies the 3.5Mb region. In *Arabidopsis*, approximately 4 million basepairs are devoted to NOR2. At this site (and similarly at NOR4 at the top of chromosome 4), the genes encoding the precursor transcripts of 18S, 5.8S and 25S ribosomal RNA are clustered. These genes are arranged head-to-tail with each transcription unit separated by an intergenic spacer (Brown and Dawid, 1969; Reeder, 1974; figure 3.14). The ~375 *Arabidopsis* rRNA genes at each NOR evolve rapidly. However, unlike most other gene families that evolve independently, every rRNA gene remains virtually identical in sequence to every other rRNA gene within the individual and even within a population, a phenomenon known as concerted evolution.

Mutations in bacterial, mitochondrial and chloroplast rRNAs resulted in mutant phenotypes and helped to elucidate the molecular mechanisms of translation. However, in all of these cases, only a limited number of rDNA copies are present (Gregory et al., 2005; Murgola et al., 1988; Shen and Fox, 1989; Thompson et al., 2001; Yassin et al., 2005). Triman et al. (1989) isolated ts 16S rRNA mutants of *Escherichia coli*, causing a disruption of

base-pairing in the secondary structure of 16S rRNA at the restrictive temperature. There was a lag period of about two generations between a shift to the restrictive temperature and cessation of growth, implying that the structural defects caused impairment of ribosome assembly. One eukaryotic organism, *Tetrahymena*, is also well-suited for the isolation of rDNA mutations (Spangler and Blackburn, 1985), since a single copy of the rDNA is present in the germline micronucleus and is amplified in the somatic macronucleus. However, in general, mutations in eukaryotic rRNA genes do not result in a mutant phenotype because of the highly repetitive nature of eukaryotic rDNA. Therefore, it is unlikely that the *dgr2* point mutation is localized in one of the rRNA repeats, thereby causing defects in protein synthesis, which would subsequently be expressed in the *dgr2* growth phenotype.

Another possibility would be that the *DGR2* gene corresponds to an essential protein-coding gene, localized in the NOR2-telomere junction region. Using telomere-specific and 18S rRNA-specific primers Copenhaver and Pikaard (1996) were able to clone the *Arabidopsis* NOR4-telomere junction. Their results clearly revealed that the NOR4 directly abuts the telomere of chromosome IV. A similar situation has been described in *Neurospora*, *Giardia* and *Tetrahymena*, where telomeric repeats abruptly join the rDNA (Adam et al., 1991; Yu and Blackburn, 1991). The phenomenon is thought to be caused by chromosome breaks that are 'healed' by telomerase, which is capable of adding consensus telomere repeats *de novo* to broken chromosome ends (Blackburn, 1991). Until now, the *Arabidopsis* NOR2-telomere junction has not been characterized. Therefore, we cannot exclude the possibility that an essential gene, corresponding to *DGR2*, could reside within this region.

A remaining possible hypothesis is that the ts phenotype of *dgr2* is caused by a point mutation in the telomere. Telomeres play an important role in cell growth and senescence and their architecture is well conserved across evolution, consisting of tandem G-rich 6bp (-GGATT- in *Arabidopsis*) repeats. The 3' terminus of the G-rich strand forms a single-strand overhang, which is used by the enzyme telomerase for the synthesis of unit repeat sequences by the addition of nucleotides with dNTPs as substrates (McEachern et al., 2000).

Because of the repetitiveness in the sequence of telomeres, individual single point mutations generally do not cause a phenotypic discrepancy (Blackburn, 1991). However, exceptions are point mutations induced at the very tip of the telomeres. This has to do with the mechanism by which telomeric length is regulated. Much data support a dynamic model, in which telomerase and the telomere act together as a homeostatic system consisting of both negative and positive regulators of telomere length. Hence, there is a continual dynamic turnover at the telomere tip, but the overall telomere length will stay constant. This makes that, even single base pair substitutions, at the very ends of telomeres can cause loss of regulation of telomeres and loss of the capability of telomeres to cap chromosomes (figure 3.18).

An example of such a mutation was the *acc* mutation in *Saccharomyces cerevisiae*, a point mutation in the binding site for Rap1 (a protein regulating telomeric length) in the telomeric repeat. This mutation caused dropping of the binding affinity by a couple of orders of magnitude (Krauskopf et al., 1996). As soon as one or two *acc* repeats were added to the ends of otherwise WT telomeres, length control was completely lost, and these telomeres became very long and highly deregulated (figure 3.18), which was reflected in a slow growth rate. This study demonstrates that, at least for yeast cells, it is possible to obtain a mutant growth phenotype by inducing point

mutations at the very ends of telomeric repeats. However, these findings cannot simply be compared with results obtained in more complex organisms, like *Arabidopsis*.

Therefore, it is not yet clear in which part of the top of chromosome II *dgr2* resides and how it thereby affects the plant phenotype.

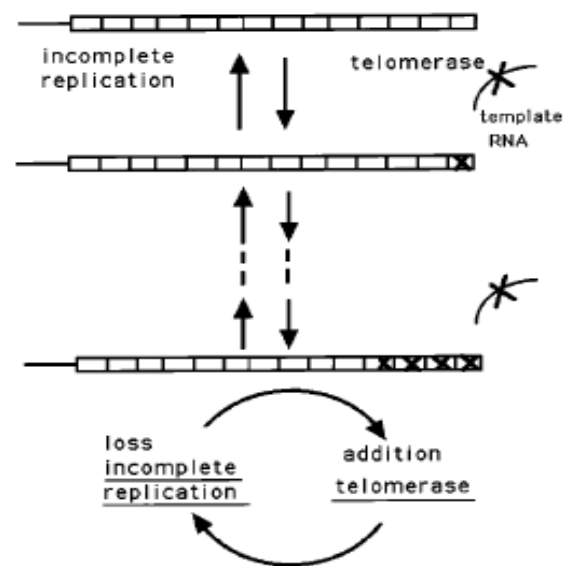


FIGURE 3.18 DYNAMIC TURNOVER AT THE TELOMERE TIP (BLACKBURN, 1991). Point mutations (X) at the very ends of telomeres can cause loss of regulation of telomeres and loss of the capability of telomeres to cap chromosomes.

3.5 CONCLUSIONS

In summary, the map-based cloning of the ts mutant *dgr1* resulted in the identification of a gene essential for the growth and development of *Arabidopsis*. *DGR1* or at2g45790 encodes eukaryotic phosphomannomutase, a protein that plays a crucial role in several fundamental cellular processes, including cell wall biosynthesis, glycoprotein synthesis and plant L-ascorbic acid production. Molecular complementation of the *dgr1* mutant with the wild type

gene is currently ongoing and will provide us with the ultimate proof that *DGR1* corresponds to the at2g45790 gene, encoding eukaryotic phosphomannomutase. In the future, *dgr1* will serve as an excellent tool in the genetic dissection of these processes. Particularly, our *dgr1* mutant will add greatly to our understanding of how plants synthesize ascorbic acid, a pathway which is still poorly understood. Only recently, it gained new attention with the isolation of a series of ozon-sensitive ascorbic acid deficient mutants, which are currently used to further unravel this metabolic pathway.

The *DGR2* gene was mapped to a chromosomal region of 3.5Mb on the top of chromosome II, consisting of NOR2 and the upper telomere. Despite the fact that these regions largely consist of non-protein coding genes, they fulfill important functions in overall plant growth and development. Since, due to the lack of informative polymorphisms, no exact map position could be assigned to *dgr2*, we have tried to characterize this mutant using an alternative approach: the molecular phenotyping of the *dgr2* mutant using gene expression analyses will be the subject of chapter four.

With the extensive genome sequence, currently available for *Arabidopsis*, phenotyping and genotyping remain the critical limiting variables in the process of map-based cloning. The genotyping problem is currently dealt with by hybridizing labeled genomic DNA to the Affymetrix *Arabidopsis* ATH1 GeneChip (Gresham et al., 2006). Using this approach several high heritability development and circadian clock traits were mapped via direct hybridization of mutant DNA to ATH1 expression arrays (Hazen et al., 2005). Direct detection of small mutations, such as an ethyl-methane

sulfonate derived single base substitutions, are still limited by array coverage and sensitivity. However, deletions for example caused by fast neutron mutagenesis are easily detected.

CHAPTER FOUR

EXPRESSION ANALYSIS OF THE TEMPERATURE-SENSITIVE *dgr2* MUTANT OF *ARABIDOPSIS THALIANA*

In the present study, we made a first attempt to characterize the temperature-sensitive growth mutant *dgr2* by means of gene expression analyses. We previously mapped the *DGR2* gene to a chromosomal region of 3.5Mb on the top of chromosome II. However, difficulties encountered during the fine-mapping of the *DGR2* gene prompted us to characterize this mutant using an alternative approach. A gene expression analysis using cDNA micro-arrays was performed, allowing us to monitor the transcriptional responses to elevated temperature in both the wild type and the *dgr2* mutant. This study revealed that a diverse set of genes involved in various fundamental cellular processes is affected by the presence of the *dgr2* mutation. We observed that genes encoding proteins involved in protein synthesis, cell wall biosynthesis and cell differentiation are primarily down-regulated in the mutant, whereas those functioning in energy metabolism (respiration), cell death and environmental stress are mainly upregulated. Further characterization of the *dgr2* mutant will be required to define the exact nature of the *DRG2* gene. Nevertheless, our results illustrate the power of expression profiling for functional gene characterization in plants, comparable to what has been previously shown by others working on yeast (Hughes et al., 2000) and tumour cells (Van 't Veer et al., 2002).

4.1 INTRODUCTION

Whole-genome expression profiling, facilitated by the development of DNA microarrays (Schena et al., 1995; Lockhart et al., 1996), represents a major advance in genome-wide functional analysis. In a single assay, the transcriptional response of each gene to a change in cellular state can be measured, whether it is disease, a process such as cell division, or a response to a chemical or genetic perturbation (Cho et al., 1998; DeRisi et al., 1997; Galitski et al., 1999; Heller et al., 1997; Holstege et al., 1998; Spellman et al., 1998). Because the relative abundance of transcripts is often tailored to specific cellular needs, most expression profiling studies have focused on the genes that respond to conditions or treatments of interest. In the same way, the cellular state can be inferred from the expression profile, and the idea that the global transcriptional response itself constitutes a detailed molecular phenotype has begun to receive attention (Eisen et al., 1998; Gray et al., 1998; Holstege et al., 1998; Marton et al., 1998; Roberts et al., 2000). Provided that the cellular transcriptional response to disruption of different steps in the same pathway is similar, and that there are sufficiently unique transcriptional responses to the perturbation of most cellular pathways, systematic characterization of novel mutants can be carried out using this strategy (Sommerville and Somerville, 1999).

In the present study, we used this approach to expand our knowledge on the temperature-sensitive (ts) *dgr2* mutant of *Arabidopsis*. To this end, we tried to map biological processes through the identification of sets of genes that showed a differential expression in the mutant relative to the wild type (WT). This way, the collection of genome-wide expression patterns, consisting of thousands of individual observations, results in a molecular phenotype or a transcriptional

fingerprint, characteristic for the mutant. During the past five years, the power of this method has been shown by several studies, for instance in yeast (Hughes et al., 2000) and tumors (Van 't Veer et al., 2002). However, the approach has never been applied for the characterization of ts mutations. Therefore this approach represents an efficient, however challenging, way to characterize unknown mutations, such as *dgr2*.

4.2 EXPERIMENTAL PROCEDURES

4.2.1 GROWTH CONDITIONS

Seeds were sterilized by a ten-minute incubation with a 1.5% sodiumhypochlorite solution (Merck, Germany). Sterile seeds were sown on MS germination medium (Murashige and Skoog, 1962) in square Petri dishes (120mm x 120mm, Greiner bio-one, Frickenhausen, Germany). After sowing, Petri dishes were placed vertically in a growth chamber (temperature specified below) under continuous light provided by V.H.O. Sylvania fluorescent white tubes at an irradiance of 50 $\mu\text{mol PAR photons m}^{-2}\text{s}^{-1}$ (Osram, Milan, Italy) or by Lumilux Cool White light tubes (Osram, Milan, Italy) at an irradiance of 50 $\mu\text{mol PAR photons m}^{-2}\text{s}^{-1}$.

4.2.2 PLANT MATERIAL

Mutant *dgr2* (Col-0) and WT (Col-0) plants were grown *in vitro* for nine days at the permissive temperature of 18°C. At day 10 mutant and WT test plants were transferred to the restrictive temperature of 28°C, whereas control plants continued to grow at 18°C.

For the microarray analysis, samples (five to ten shoots per sample) were taken at six consecutive time points

after the temperature shift (0hrs, 12hrs, 36 hrs, 60hrs, 84hrs, 96 hrs). Three days after the temperature-shift (at time point 84hrs) test plants were rescued by transferring them back to the permissive temperature. A pool of WT control plants served as a reference. A biological repeat was performed, which served at the same time as a color swap.

For the cDNA-AFLP analysis, samples (five to ten shoots per sample) were taken at nine consecutive time points after the temperature shift (0hrs, 0.25hrs, 0.5hrs, 1hr, 2hrs, 4hrs, 6hrs, 8hrs, 10hrs). No rescue step was incorporated. Both a pool of WT (ten shoots) and a pool of mutant (ten shoots) control plants served as a reference.

4.2.3 RNA EXTRACTION AND cDNA SYNTHESIS

Total RNA was prepared from each sample using the LiCl precipitation protocol (Sambrook, 1988).

First strand cDNA synthesis was done by reverse transcription using a biotinylated oligo-dT₂₅ primer (Genset) and SuperscriptII (Gibco-BRL). Second strand synthesis was performed by strand displacement with *E. coli* ligase (Gibco-BRL), DNA polymerase I (USB) and Rnase-H (USB). Further cleaning of the cDNA occurred by elution through DNA Spin columns (Amersham Pharmacia Biotech).

4.2.4 cDNA AFLP ANALYSIS

Double stranded cDNA (500 ng) was used for AFLP analysis based on the method described by Vos et al. (1995) and Bachem et al. (1996) with the below described modifications. The restriction enzymes used were *Bst*YI and *Mse*I (New England Biolabs) and the digestion was done in two separate steps. After the first restriction digest with *Bst*YI, the 3' end fragments were collected on Dynabeads

(Dynal) by means of their biotinylated tail, while the other fragments were washed away. After digestion with *Mse*I, the released restriction fragments were collected and used as templates in the subsequent AFLP steps. The adapters and primers used are:

<i>Bst</i> YI adapter:	5' – CTCGTAGACTGCGTAGT – 3'
	5' – GATCACTACGCAGTCTAC – 3'
<i>Mse</i> I adapter:	5' – GACGATGAGTCCTGAG – 3'
	5' – TACTCAGGACTCAT – 3'
<i>Bst</i> YI-primer:	5' – GACTGCGTAGTGATC(T/C) _{N1-2} – 3'
<i>Mse</i> I-primer :	5' – GATGAGTCCTGAGTAAN ₁₋₂ – 3'

For pre-amplifications, an *Mse*I-primer without selective nucleotides was combined with a *Bst*YI-primer containing either a T or a C as 3' most nucleotide. PCR conditions were applied as described (Vos et al., 1995). The obtained amplification mixtures were diluted 600-fold and 5 µL was used for selective amplifications using a P³³-labeled *Bst*YI-primer and the AmpliTaq-Gold polymerase (Roche) following the described procedure (Vos et al., 1995). Amplification products were separated on 5% polyacrylamide gels using the Sequigel system (Biorad). Sodium acetate (22 mg/L) was added to the buffer in the bottom tank to avoid run-off of the smallest fragments. Dried gels were exposed to Kodak Biomax films and in an additional approach scanned in a Phosphor-Imager apparatus (Molecular Dynamics).

Bands corresponding to differentially expressed transcripts were cut out from the gel, the DNA was eluted in TE buffer and re-amplified using the same conditions as for selective amplification.

Sequence information was obtained by direct sequencing of the re-amplified PCR product using the selective *Bst*YI-primer. The obtained sequences were compared against nucleotide and protein sequences present in the publicly available databases by BLAST sequence alignments (Altschul et al., 1997).

Scanned Molecular Dynamics gel images were quantitatively analyzed using AFLP-QuantarPro™ (Keygene, The Netherlands). This software has been

designed and optimized for accurate lane definition, fragment detection and quantification of band intensities. All visible AFLP fragments were scored and individual band intensities were measured per lane. The obtained data were used to determine the quantitative expression profile of each transcript. Raw data were corrected for differences in total lane intensities, which may occur due to loading errors or differences in the efficiency of PCR amplification with a given primer combination (PC) for one or more time points. The correction factors were calculated based on constant bands throughout the time course. For each PC, a minimum of ten invariable bands was selected and the intensity values were summed per lane. Each of the summed values was divided by the maximal summed value to give the correction factors. Finally, raw data generated by AFLP-QuantarPro were divided by the correction factors.

4.2.5 MICROARRAY ANALYSIS

Microarray hybridizations were performed at the MAF (MicroArray Facility, VIB, Leuven; <http://www.microarray.be/>). We used the VIB 6K *Arabidopsis* microarray, spotted with 6008 *Arabidopsis* PCR-amplified cDNA-clones or EST's with a size ranging from 0.5 to 2.5 kb. Each clone was spotted twice.

Test samples of both the WT and the mutant were labeled with the green fluorescent dye Cy3, whereas the reference was labeled with the red fluorescent dye Cy5. According to a reference design each test sample was then hybridized with the reference in a one to one ratio on a slide. Finally, Cy3 and Cy5 intensities were measured for each spot on the slide. A biological repeat was performed and served at the same time as a color swap (Cy5=test en Cy3=control). This resulted for each time point in four observations per genotype.

For the statistical analysis, spot ratios between Cy3- and Cy5-values were calculated. These ratios were subsequently Lowess-fit normalized (Yang et al., 2002). This normalization step corrects the systematically higher Cy3-intensities, especially occurring at low total spot intensities.

Genes with a statistically significant positive expression signal at more than half of the time points in both the mutant and the WT were selected. For each array, the cut-off (95-percentile) for a positive signal was calculated based on 96 (Cy3- and Cy5-) values of 24 *Bacillus* clones, spotted at different positions on the 6K array. 4797 genes (80%) met to this criterium and were further analyzed. Subsequently, a sequential ANOVA (Wolfinger et al., 2001) was performed, which accounts for experiment-wide systematic effects (e.g. array effects) and extracts appropriate information about the treatment effects, i.e. temperature, genotype, and the interaction which is of primary interest for each gene. We tested for significant differences between these effects by using the Wald-statistic, which is comparable with an F-statistic.

Cluster analysis of the expression data was performed using two methods. In the first method, a hierarchical, average linkage clustering was performed using the Cluster software developed by Eisen and co-workers (1998). For the visualization of the results, the Treeview software (Eisen et al., 1998) was applied. Both software packages can be downloaded for free at <http://rana.lbl.gov/EisenSoftware.html>. The resulting cluster shows a hierarchical tree, in which each line represents a gene. By clicking on a node, the corresponding cluster can be viewed in more detail. The expression values are presented with color codes for each measured time point. The second method is Adaptive Quality Based Clustering (De Smet et al., 2002) and is accessible for use on <http://www.esat.kuleuven.ac.be/~thijs/Work/Clustering.html>. As an output, genes are grouped in different clusters and are presented in graphs, where their average

expression pattern is shown in function of time. Both methods require the input expression data to be submitted in a text format (.txt), where a given row indicates data for one single gene and each column represents an expression value for a time point.

4.3 RESULTS

To expand our knowledge about the biological processes that are affected by the *ts dgr2* mutation, we performed a gene expression analysis using cDNA microarrays. This study allowed us to identify the transcriptional response to high temperature in the WT and the *dgr2* mutation.

4.3.1 MICROARRAY ANALYSIS (1): TRANSCRIPTIONAL RESPONSE TO ELEVATED TEMPERATURES IN WILD TYPE PLANTS

As a first step in the analysis of the microarray data we specifically focussed on the effect of the temperature treatment. Samples of WT plants (five to ten shoots per sample) were taken at six consecutive time points after the temperature shift (0hrs, 12hrs, 36 hrs, 60hrs, 84hrs, 96 hrs). Three days after the temperature-shift (at time point 84hrs) test plants were rescued by transferring them back to the permissive temperature. A pool of WT control plants served as a reference. A biological repeat was performed, which served at the same time as a color swap. We submitted the WT dataset to an ANOVA (analysis of variance) to capture the effects of the temperature shift on gene expression. As a result, 1098 genes were differentially expressed ($p < 0.001$).

The selected genes were categorized by function according to the classification used by MatDB (MIPS *Arabidopsis thaliana* database). Additional information regarding

(potential) gene function was gathered from the literature. A function could be assigned to approximately 40% of the genes (422 genes).

Among these 422 differentially expressed genes all major functional categories were represented (figure 4.1). Almost 50% of the stress and defense genes represented on the microarray were differentially expressed upon the temperature shift. Furthermore, 36% of the genes involved in carbohydrate metabolism, 30% of those functioning in amino acid metabolism and 32% playing a role in energy metabolism were significantly up- or downregulated as a consequence of the temperature changes. The remaining functional classes were less represented and each comprised between 12-18% of the genes present on the array.

We clustered the collection of 1098 differentially regulated genes based on the gene expression profiles using hierarchical (Eisen et al., 1998) and adaptive quality based clustering (De Smet et al., 2002) (figure 4.2). Both clustering methodologies resulted in three distinct groups of genes, characterized by the transcriptional response to the temperature up- and/or downshift (=rescue-step).

A first group of genes shows a rapid and immediate transcriptional response to both temperature changes. Within this group, we were able to distinguish a first subcluster showing induced and a second subcluster showing repressed gene expression at the higher temperature. The first subcluster consists of 59 genes, which are all stress- or defense related. The second subcluster comprises approximately 200 genes belonging to various functional classes, although a vast majority plays a significant role in transcription and energy metabolism.

A second group of genes shows a gradual transcriptional response to the temperature upshift. Again two subclusters were identified, characterized by induced and repressed gene expression,

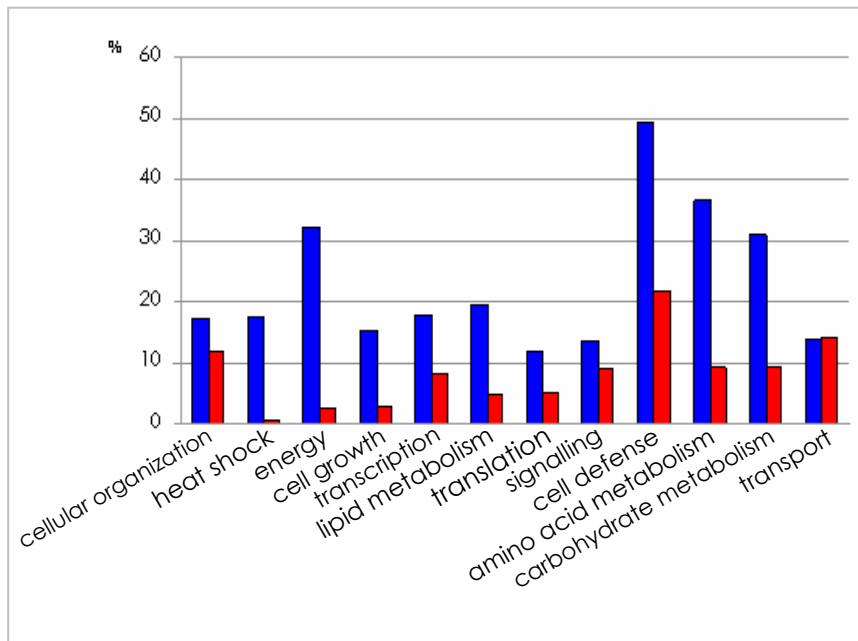


FIGURE 4.1 FUNCTIONAL CLASSIFICATION OF THE GENES THAT WERE DIFFERENTIALLY EXPRESSED (422) UPON TEMPERATURE INCREASE IN WT PLANTS IN A MICROARRAY EXPERIMENT. For each functional class, blue bars indicate the percentage of genes represented on the microarray that are differentially expressed upon temperature elevation. For each functional class, red bars represent the percentage of *Arabidopsis* genes represented on the microarray. Almost 50% of the stress and defense genes represented on the microarray were differentially expressed upon the temperature shift. Furthermore, 36% of the genes involved in carbohydrate metabolism, 30% of those functioning in amino acid metabolism and 32% playing a role in energy metabolism were significantly up- or downregulated as a consequence of the temperature changes. The remaining functional classes were less represented and each comprised between 12-18% of the genes present on the array.

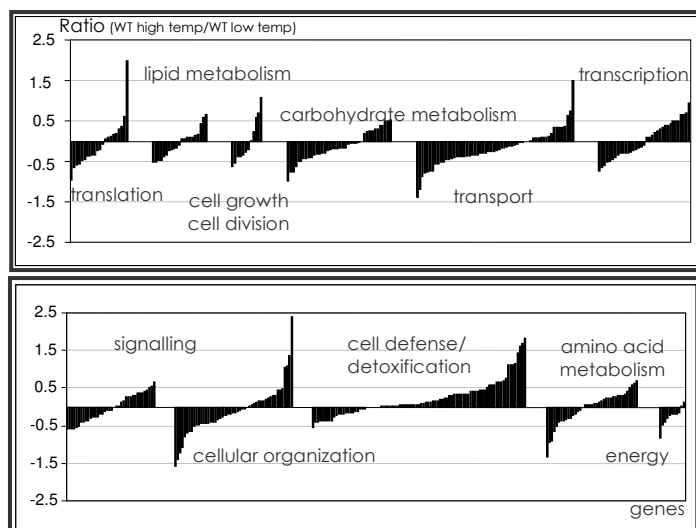


FIGURE 4.3 A PLOT OF THE UP- AND DOWN REGULATION OF GENES IN RESPONSE TO A TEMPERATURE SHIFT IN WT PLANTS, CLASSIFIED ACCORDING TO FUNCTION. Gene expression ratios, obtained at a representative time point (60 hrs after the temperature shift), were plotted and sorted by gene function. The category of detoxification and cell defense is mainly characterized by induced gene expression, whereas transport, energy, and carbohydrate metabolism mainly consist of down-regulated genes. Each individual gene is represented by a vertical black bar.

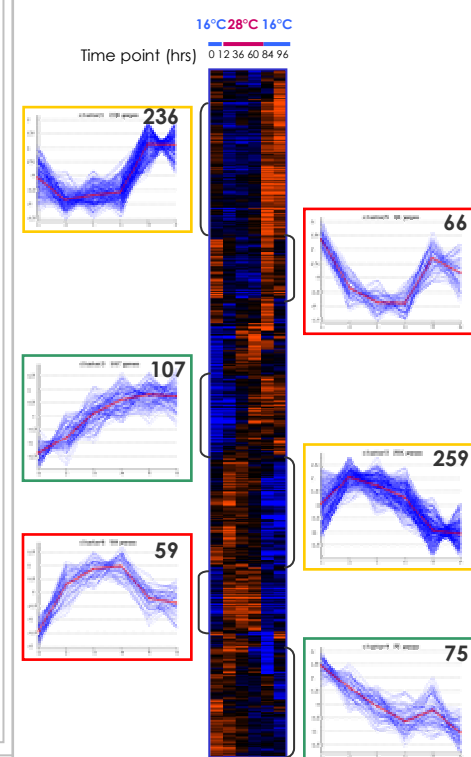


FIGURE 4.2 HIERARCHICAL AND ADAPTIVE QUALITY BASED CLUSTERING OF EXPRESSION PROFILES MODULATED BY HIGH TEMPERATURE IN WT PLANTS. Both clustering methodologies resulted in three distinct groups of genes, characterized by the transcriptional response to the temperature up- and/or downshift. A first group of genes (—) shows a rapid and immediate transcriptional response to both temperature changes. Within this group, we were able to distinguish a first subcluster showing induced and a second subcluster showing repressed gene expression at the higher temperature. A second group of genes (—) shows a gradual transcriptional response to the temperature upshift. Again two subclusters were identified, characterized by induced and repressed gene expression, respectively. A third group of genes (—) shows a rapid transcriptional response (exclusively) to the temperature downshift. Expression is downregulated in the first and upregulated in the second subcluster.

■ Induction
■ No differential gene expression
■ Repression
■ Missing value

respectively. Both subclusters contain genes acting in a diversity of biological processes.

A third group of genes shows a rapid transcriptional response (exclusively) to the temperature downshift. Expression is downregulated in the first and upregulated in the second subcluster, with the latter consisting of a significant number of cold responsive genes.

Finally, to assess whether up- or down regulation of genes was specific for particular functional classes, we plotted the gene expression ratios, obtained at a representative time point (60 hrs after the temperature shift), and sorted these by the function of the corresponding genes (figure 4.3). The category of detoxification and cell defense is mainly characterized by induced gene expression, whereas transport, energy, and carbohydrate metabolism consist of a vast number of down-regulated genes. Other classes show a similar degree of induced and repressed gene expression.

4.3.2 MICROARRAY ANALYSIS (2): TRANSCRIPTIONAL RESPONSE TO *dgr2* IN MUTANT PLANTS

In a second microarray analysis, we made a first attempt to map the biological processes, specifically affected by the *dgr2* mutation. Mutant samples (five to ten shoots per sample) were taken at six consecutive time points after the temperature shift (0hrs, 12hrs, 36 hrs, 60hrs, 84hrs, 96 hrs). Three days after the temperature-shift (at time point 84hrs) test plants were rescued by transferring them back to the permissive temperature. A pool of WT control plants served as a reference. A biological repeat was performed, which served at the same time as a color swap. To this end, we identified sets of genes that were differentially expressed in the mutant relative to the WT. This way, the collection of genome-wide expression patterns results in a molecular

phenotype or a transcriptional fingerprint, characteristic for the mutant.

We selected 463 (approximately 8%) genes differentially expressed ($p < 0.001$) in the mutant compared to the WT. Again, a functional classification was performed according to the MatDB. An accurate function could be assigned to approximately 35% (161 genes) of the selected genes.

To determine which functional classes were significantly over- or underrepresented in our collection of 161 genes, a χ^2 -test was performed. For none of the functional classes statistically significant χ^2 -values were obtained. Therefore, we tried to assign a more specific function to each gene by subdividing each general category in several subcategories. Seven subclasses appeared to be significantly ($p < 0.05$) over- or underrepresented: intracellular transport, respiration, cell differentiation, peptide transporters, environmental stress, protein synthesis and RNA localization (table 4.1). A broad diversity of processes appear to be significantly affected by the mutation. After calculating the percentages of induced and repressed gene expression, we observed that the categories protein synthesis, cell wall biosynthesis and cell differentiation primarily comprised down-regulated genes, whereas all the other classes mainly consisted of upregulated genes.

4.3.3 EARLY TRANSCRIPTIONAL RESPONSE TO *dgr2* IN MUTANT PLANTS DETECTED BY cDNA AFLP ANALYSIS

In order to identify very early effects of the mutation on gene expression, we performed a limited cDNA AFLP analysis. Samples were taken at nine consecutive time points within a time course of 12 hours after the temperature shift.

132 differentially expressed cDNA fragments were visually selected and 106 of these could be efficiently sequenced.

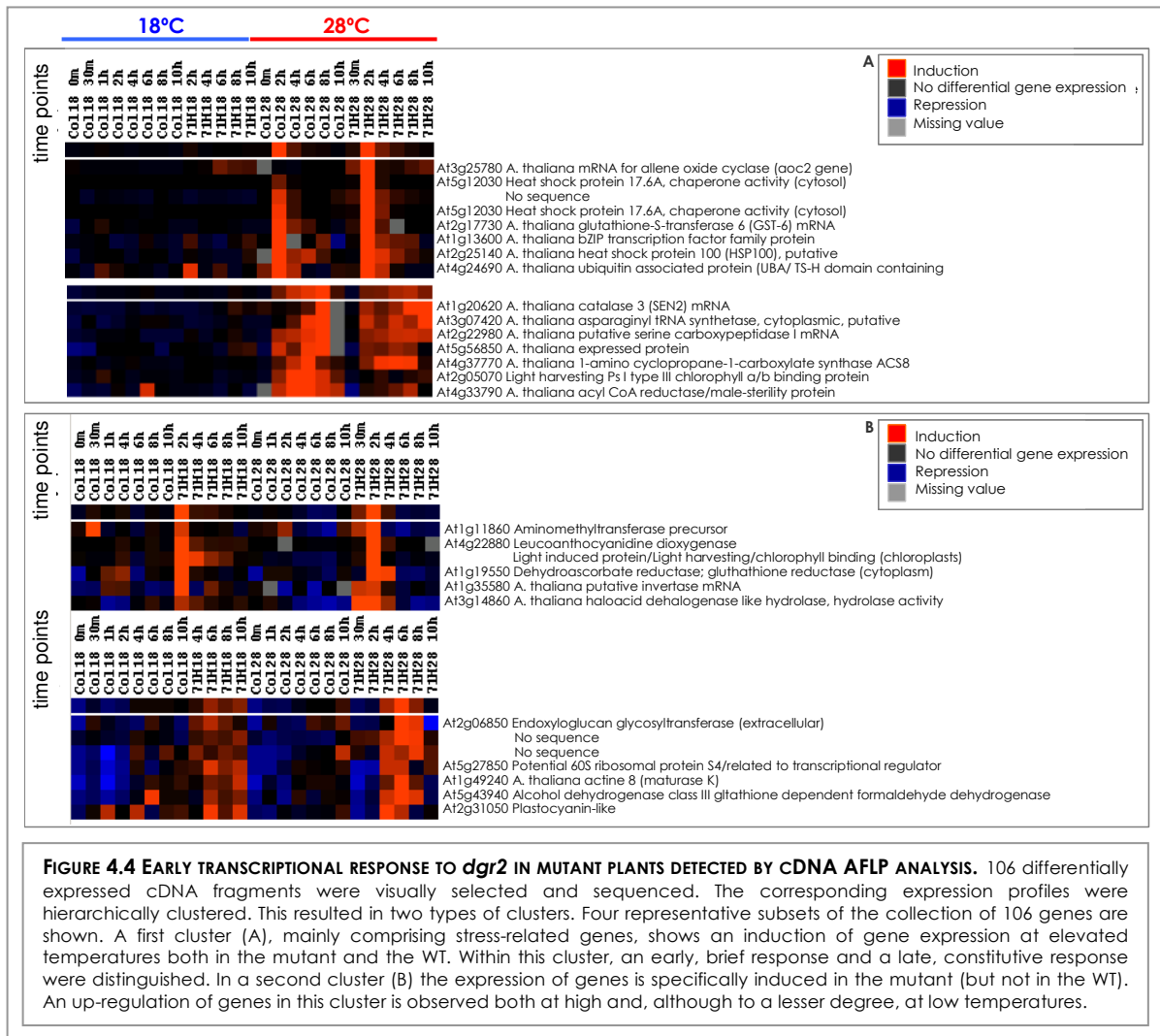


TABLE 4.1 FUNCTIONAL CLASSES OF GENES, FOUND TO BE SIGNIFICANTLY AFFECTED BY THE *dgr2* MUTATION AFTER MICROARRAY GENE EXPRESSION ANALYSIS.

FUNCTIONAL CLASS	NUMBER OF GENES	$\chi^2_{(0.05;1)}$ -VALUE	UP-/DOWN-REGULATION
Cell death	4	3.45	Up
Cell wall biosynthesis	15	3.15	Down
mRNA transcription	51	3.82	Up
Proteolytic degradation	15	3.33	Up
Intracellular transport	13	7.19	Up
Respiration	15	5.32	Up
Cell differentiation	15	4.96	Down
Peptide transporters	4	5.18	Up
Environmental stress	20	5.46	Up
Protein synthesis	6	8.10	Down
RNA localization	8	4.70	Up

$\chi^2_{(0.05;1)} = 3.84$

The obtained sequences were compared against nucleotide and protein sequences present in the publicly available databases by BLAST sequence alignments.

The corresponding expression profiles were hierarchically clustered (Eisen et al., 1998). This resulted in two types of clusters (figure 4.4).

A first cluster, mainly comprising stress-related genes, shows an induction of gene expression at elevated temperatures both in the mutant and the WT. Within this cluster, an early, brief response and a late, constitutive response were distinguished. In a second cluster the expression of genes is specifically induced in the mutant (but not in the WT). An up-regulation of genes in this cluster is observed both at high and, although to a lesser degree, at low temperatures. This is in agreement with what we expect when dealing with a *ts* mutation: the molecular deficiencies are moderately expressed at the permissive temperature, whereas at the restrictive temperature severe molecular defects are induced, giving rise to a visible mutant phenotype. Also in this cluster, an early and a late, constitutive response are distinguished. The first subcluster consists primarily of stress related genes, whereas the second comprises genes involved in transcription and cell wall biosynthesis (figure 4.4).

4.4 DISCUSSION

In the present study, we made an attempt to map the biological processes that are affected by the *ts dgr2* mutation. A genome-wide expression analysis using cDNA microarrays was performed, which allowed us to identify the transcriptional responses to both the elevated temperatures in the WT and to the *dgr2* mutation in the mutant.

4.4.1 GENOME-WIDE TRANSCRIPTIONAL RESPONSE TO ELEVATED TEMPERATURES IN WILD TYPE *ARABIDOPSIS* PLANTS

1098 of the 6008 genes present on the microarray were differentially expressed in WT plants upon temperature shift. In this collection, genes from all the major functional categories are represented. The large number of stress regulated transcriptome changes observed in our study underscores the difficulty of understanding the global context of a particular stress response. Extrapolating this to the entire genome of *Arabidopsis*, approximately 5000 genes are potentially affected in their expression upon higher temperature. Because all aspects of plant physiology are impacted by stress, we consider a major change in the transcriptome to be reasonable. Genes encoding proteins involved in stress/defense were generally upregulated and comprised 22% (91 genes) of the collection of 422 genes to which a function could be assigned. This category included genes encoding proteins of the antioxidant response (glutathione-S-transferases, peroxidases), the SAR response (pathogenesis-related genes) and cell rescue (heat shock proteins, heat shock transcription factors). Genes that are reported to be induced in response to other stresses, such as drought (Gosti et al., 1995), and dehydration (Kiyosue et al., 1993; Koizumi et al., 1993; Nylander et al., 2001) were also identified in our stress collection of differentially expressed genes, underscoring the overlap between the different stress responses.

Furthermore, genes involved in carbohydrate metabolism are highly represented in our collection of differentially expressed genes (9% or 39 genes). Soluble sugars, especially sucrose, glucose, and fructose, play an obviously central role in plant structure and metabolism at the cellular and whole-organism levels. They are involved in the responses to a number of stresses, and they act as nutrient and metabolite

signalling molecules that activate specific or hormone-crosstalk transduction pathways, thus resulting in important modifications of gene expression patterns. Various metabolic reactions and regulations directly link soluble sugars with the production rates of reactive oxygen species, such as mitochondrial respiration or photosynthesis regulation, and, conversely, with anti-oxidative processes, such as the oxidative pentose-phosphate pathway and carotenoid biosynthesis. Moreover, stress situations where soluble sugars are involved, such as heat shock, chilling, herbicide injury, or pathogen attack, are related to important changes in reactive oxygen species balance. These converging or antagonistic relationships between soluble sugars, reactive oxygen species production, and anti-oxidant processes are generally confirmed by current transcriptome analyses, and suggest that sugar signalling and sugar-modulated gene expression are related to the control of oxidative stress (Coue   et al., 2006; Sulmon et al., 2006).

Genes encoding proteins involved in amino acid metabolism constituted (together with carbohydrate metabolism) the second largest group (9% or 39 genes) in the collection. A majority of the genes in this category are involved in proteolysis. Proteolysis of important regulatory proteins is a key aspect of cellular regulation in eukaryotes (del Pozo et al., 1999) and there is evidence that the ubiquitin-proteasome pathway is important in implementation of the plant defense response (Azevedo et al., 2002; Austin et al., 2002; Liu et al., 2002). Proteasome subunit genes are induced in response to stresses and some regulatory subunits were identified in the collection, such as RING finger proteins, which are key components of the ubiquitin-proteasome pathway (Saurin et al., 1996). The RING finger motif is thought to mediate protein-protein interactions and E3 ligase complex assembly. Certain RING finger proteins are rapidly induced by elicitors in *Arabidopsis* and may be involved in the rapid degradation of regulatory proteins during

early stages of pathogen attack (Salinas-Mondragon et al., 1999). Four different genes encoding RING/RING-H2-finger proteins were identified in our collection.

In summary, the collection of high temperature-induced genes comprises a broad repertoire of stress-responsive genes encoding proteins that are involved in both the initial and subsequent stages of the physiological response to stress.

4.4.2 GENOME-WIDE TRANSCRIPTIONAL RESPONSE TO THE TEMPERATURE-SENSITIVE *dgr2* MUTATION

In a second analysis, we identified genes that showed a differential expression in the *dgr2* mutant relative to the WT. This way, the collection of expression patterns results in a molecular phenotype or a transcriptional fingerprint, characteristic for the mutant. In the mutant 463 genes (~8%) were differentially expressed ($p < 0.001$).

A very diverse set of fundamental cellular processes is affected by the presence of the *dgr2* mutation. Genes involved in protein synthesis, cell wall biosynthesis and cell differentiation are primarily down-regulated, whereas the categories of energy metabolism (respiration), cell death and environmental stress mainly comprise upregulated genes. The large number and diversity of these cellular processes might suggest that the observed gene expression profile corresponds to a secondary physiological response to the mutation, which makes it very difficult to pinpoint the processes that are primarily affected by the mutation.

Our previous results outlined in chapter three show that *dgr2* maps to a 3.5 Mb region on the upper part of chromosome II. We postulated different hypotheses concerning the location of the gene in the NOR region, in the NOR-telomere junction or in the upper telomere of chromosome II. In higher eukaryotes, no rRNA mutants have been isolated (or functionally characterized) until now due

to the high copy number of rDNA in these species. Furthermore, in most eukaryotic species, the NOR-telomere junction region has been poorly characterized. However, telomere-deficient mutants in yeast (*Saccharomyces cerevisiae*) have been described by Nautiyal et al. (2002). They characterized these mutants by monitoring the genome-wide transcriptional response, whereby sets of genes were defined that are involved in cellular responses to altered telomere structure or function. When comparing our expression data from the *dgr2* mutant with the results described in this report, some coincidence is observed. Nautiyal et al. report a marked up-regulation of energy production genes, more specifically involved in oxidative phosphorylation. Furthermore, they observed an up-regulation of approximately 50% of the genes involved in the environmental stress response. Finally, a vast majority of the genes involved in protein synthesis was reported to be down-regulated.

These results are similar to what we have observed in the *dgr2* mutant. Additionally, genes involved cell differentiation and cell wall biosynthesis are down-regulated in *dgr2*, which is also known to be the case when telomere function is disturbed and which eventually results in senescence and cell death. The latter process was also significantly represented in the collection of differentially expressed genes in *dgr2*.

Despite the similarities between the transcriptional responses of the *dgr2* mutant and the yeast telomere-deficient mutants, we have to take caution making this comparison since the observations were made in different systems and under different conditions.

4.4.3 EARLY EFFECTS OF THE *dgr2* MUTATION ON THE TRANSCRIPTIONAL RESPONSE IN THE MUTANT

A limited cDNA-AFLP analysis allowed to identify induced gene

expression in the mutant as early as 30 minutes after the temperature shift.

Based on the identity of the selected differentially expressed genes, we observed that a very early and a late general stress response are induced upon temperature increase in both the mutant and the WT. This can be considered as the onset of the general response to heat stress. Furthermore, mutant-specific stress-related genes are induced, which reveals that the mutant is more stress-sensitive than the WT. Further conclusions concerning gene expression patterns based on this limited study would not be accurate and should be based on a more detailed genome-wide analysis, using the early time points that we have defined as informative. This way fewer secondary effects will be encountered.

4.5 CONCLUSION

We applied genome-wide expression profiling to identify the biological processes affected by the *dgr2* mutation, and hence, indirectly obtain a vision on the potential function of the *DGR2* gene. The results we obtained do indeed show that several crucial cellular processes are affected by mutating the *DRG2* gene. It is the first time that this approach has been applied to the characterization of temperature-sensitive mutations. Whilst the expression data did not allow us yet to assign a clear-cut function to the *DGR2* gene, our results already give an indication of the power of expression profiling for functional gene characterization, comparable to what has been previously shown by others working on yeast or tumor cells. Further characterization of the *drg2* mutant will be required to define the exact nature of the *DRG2* gene and to pinpoint the processes that are primarily affected by the *dgr2* mutation.

CHAPTER FIVE

CONCLUSIONS AND PERSPECTIVES

CONCLUSIONS AND PERSPECTIVES

In this work we have studied a collection of 12 newly isolated temperature-sensitive *Arabidopsis* mutants and performed a detailed molecular characterization of two of these mutants, named *dgr1* and *dgr2*. These two mutants, which are both defective in overall growth at the restrictive temperature, were chosen to test the feasibility of using map-based cloning to identify novel genes essential for the growth and development of the plant. Whereas *dgr1* was selected for its irreversible defective phenotype, *dgr2* was chosen as a representative of a mutant subclass with a phenotype that is reversible when temperature is switched back to normal levels.

We were successful in fine-mapping and cloning the *DGR1* gene, and have provided evidence that it encodes a eukaryotic phosphomannomutase, an enzyme that had not yet been functionally identified in *Arabidopsis*. Homologues of this enzyme in other eukaryotic species have been shown to play a crucial role in several fundamental cellular processes, including in cell wall biosynthesis and glycoprotein synthesis. The plant enzyme is thought to be involved in L-ascorbic acid production. Molecular complementation of the *dgr1* mutant with the wild type gene is currently ongoing and should provide the ultimate proof that *DGR1* corresponds to the *at2g45790* gene, encoding eukaryotic phosphomannomutase. In the future, functional enzymatic characterization should allow to assign a role to *at2g45790* in cell wall biosynthesis, vitamin C production and/or protein glycosylation. If this is the case, *dgr1* will serve as an excellent tool in the genetic dissection of these processes. Particularly, our *dgr1* mutant will add greatly to our understanding of how plants synthesize ascorbic acid, a pathway which is still poorly understood. Only recently, it gained new attention with the isolation of a series of ozon-sensitive ascorbic acid deficient mutants, which are currently used to further unravel this metabolic pathway.

The *DGR2* gene mapped to a chromosomal region of 3.5Mb on the top of chromosome II. This 3.5Mb region comprises one of the few segments of the *Arabidopsis* genome that has not been sequenced because it consists mainly of rDNA repeats, the NOR region. Because of the lack of sequence information there are no molecular markers available in this region, which hampered further fine-mapping of the *DGR2* gene. Difficulties encountered in the mapping of the *DGR2* gene prompted us to explore an alternative approach to characterize the *dgr2* mutant. We applied genome expression profiling using cDNA micro-arrays to identify the biological processes affected by the *dgr2* mutation, and hence, indirectly obtain a clue of the potential function of the *DGR2* gene. A genome-wide expression analysis using cDNA micro-arrays was performed comparing the transcriptional responses to elevated temperatures in both wild type and the *dgr2* mutant. The results we obtained do indeed show that a diverse set of genes involved in various fundamental cellular processes is affected by the *dgr2* mutation. Since the genome-wide expression data did not allow us yet to assign a clear-cut function to the *DGR2* gene, further characterization of the *dgr2* mutant will be required to define the exact nature of the *DRG2* gene and to pinpoint the processes that are primarily affected by the *dgr2* mutation.

Ultimately, to be able to fully characterize *dgr2* gene, the underlying mutation would have to be identified by DNA sequencing. The ~3.5Mb region on the top of chromosome II, to which the *dgr2* mutation was mapped, comprises the non-sequenced NOR2 region, the NOR2-telomere junction and the upper telomere. Sequencing the NOR2 region would not

only be tremendously time-consuming and expensive, but also unlikely to succeed because it would require finding single base differences between wild type and mutant within a 3.5 Mb segment, which is technically not feasible. Alternatively, the TILLING approach (Targeting Induced Local Lesions in Genomes) might provide a way to determine whether a point mutation is present in this area or not. TILLING is a method that allows a directed identification of mutations in a specific gene or nucleotide sequence, and could be used to search for differences between wild type and mutant. Also this approach may not succeed because of the highly repetitive nature of the sequence. Furthermore it is difficult to imagine how a mutation in one of the ~300 ribosomal DNA repeats could lead to a temperature-sensitive phenotype. Consequently, the most likely hypothesis is that the *dgr2* mutation lies in either in the NOR2-telomere junction region or the telomere itself. Until now, there is no experimental evidence for the presence of protein coding genes in NOR-telomere junction regions. However, to exclude this possibility a fragment bridging the NOR2 and the telomere could be sequenced using ribosomal repeat and a telomeric repeat specific primer. The upper telomere, being ~5kb in length, could be sequenced using telomere specific primers, designed based on the sequence of the telomeric repeats.

Our work has illustrated that the identification of novel essential genes can be carried out using the classical forward genetics approach of characterizing temperature-sensitive mutants. Especially for organisms such as *Arabidopsis*, for which sufficient genetic resources including high density genetic maps and a wealth of DNA polymorphisms are available, map-based cloning of genes is nowadays a relatively straightforward procedure. Nonetheless, map-based cloning of randomly induced mutations remains very laborious and time consuming, and as we experienced with *dgr2*, does not always succeed. The principal rate-limiting steps in map-based cloning are the requirement for (multiple) back-crosses, the generation of large mapping populations, and ultimately the sequencing of the candidate genes. Consequently, despite the tremendous technical progress, map-based cloning is not yet a high-throughput approach that can be used to characterize large numbers of genes. With the elucidation of the *Arabidopsis* genome sequence and the development of novel reverse genetic tools, one of the ultimate goals of *Arabidopsis* research is the systematic characterization of the function of all genes in a high-throughput manner. In this perspective, the high throughput identification of genes that are essential for growth and development remains an important target. For this, the recently developed inducible or conditional reverse genetic systems, such as conditional RNAi induced inhibition of gene expression, could provide a more efficient way to identify novel essential genes in *Arabidopsis* in the future.

CHAPTER SIX

SAMENVATTING

SAMENVATTING

Gedurende dit onderzoek hebben we een collectie van 12 nieuwe temperatuursensitieve *Arabidopsis* mutanten geïsoleerd. Twee van deze mutanten, namelijk *dgr1* en *dgr2*, werden geselecteerd voor een verdere gedetailleerde moleculaire karakterisering. Deze twee mutanten, beide verstoord in hun algemene groei en ontwikkeling bij de restrictieve temperatuur, waren een ideaal onderzoeksmiddel om nieuwe genen te identificeren die essentieel zijn voor groei en ontwikkeling van een gehele gedifferentieerde plant. *Dgr1* werd specifiek geselecteerd voor zijn irreversibel fenotype, terwijl *dgr2* werd gekozen als representatief exemplaar voor een subklasse van mutanten die alle een reversibel groeifenotype vertoonden wanneer ze terug naar een normale temperatuur werden overgeplaatst.

We zijn erin geslaagd het *DGR1* gen te karteren en vonden dat het gen codeert voor een eukaryotisch fosfomannomutase. Het is de eerste keer dat dit enzyme functioneel werd gekarakteriseerd in *Arabidopsis*. Homologen van dit enzyme in andere eukaryoten waren gekend en spelen een cruciale rol in diverse fundamentele cellulaire processen, zoals de celwandbiosynthese, glycoproteïnesynthese en vitamine C productie in planten.

Het *DGR2* gen werd gekarteerd tot een chromosomaal segment van 3.5 Mb op de top van chromosoom II. Dit gebied bevatte slechts een zeer klein gesequeneerd 38.5 kb fragment voorafgaand aan de NOR regio en het telomeer van chromosoom II, wat verdere fijnkartering van het *DGR2* gen onmogelijk maakte. Daarom poogden we om deze mutant te karakteriseren, gebruik makende van een alternatieve strategie. We voerden een genoomwijde expressie analyse met cDNA microarrays uit, wat ons toeliet de globale transcriptionele respons tengevolge van temperatuursverhoging te observeren in het wild type en in de *dgr2* mutant. Deze studie toonde aan dat een diverse set van genen, betrokken in allerlei fundamentele cellulaire processen, werd geaffecteerd door de aanwezigheid van de *dgr2* mutatie. Deze genexpressie-data op zich lieten ons nog niet toe om de specifieke biologische processen die in eerste instantie zijn geaffecteerd door de *dgr2* mutatie, te capteren. Verdere (genoomwijde) karakterisering van de *dgr2* mutant is dan ook vereist om de exacte aard van het gemuteerde gen te definiëren.

Ons onderzoek benadrukt de voordelen van het combineren van methoden uit de klassieke genetica en uit het moderne genomics tijdperk. Traditionele technieken, zoals het screenen van EMS mutanten voor specifieke fenotypes, waren vroeger reeds gekend als succesvolle methoden om genen te identificeren, betrokken in fundamentele processen van ontwikkeling. Verder heeft het gebruik van temperatuursensitieve mutanten ons toegelaten om planten, die ernstig verstoord waren in groei en ontwikkeling, op te pikken en te bestuderen. Door gebruik te maken van klassieke positionele klonering, een methode die in *Arabidopsis* aanzienlijk werd vergemakkelijkt door de beschikbaarheid van grote hoeveelheden aan sekwentie-informatie en genetische kaarten met hoge resolutie, kon het *DGR1* gen worden gekloneerd. Door de moeilijkheden bij het karteren van het *DGR2* gen, werden we gedwongen een alternatieve strategie te exploreren om de *dgr2* mutant te karakteriseren. We voerden een genoomwijde expressie-analyse uit gebruik makende van cDNA microarrays om de biologische processen die geaffecteerd werden door de *dgr2* mutatie te identificeren en zodanig onrechtstreeks een indicatie te verkrijgen over de mogelijke functie(s) van het *DGR2* gen. Onze resultaten tonen inderdaad aan dat verschillende essentiële processen worden beïnvloed in de *dgr2* mutant bij hoge temperatuur. Het is de eerste keer dat deze aanpak wordt gebruikt voor het karakteriseren van relatief complexe temperatuursensitieve mutaties. Desalniettemin illustreren onze

resultaten reeds de kracht van expressie-analyse om mutanten in planten te karakteriseren op moleculair niveau, vergelijkbaar met wat reeds aangetoond werd door anderen in diverse organismen.

REFERENCES

- A**arts, M.G.M., Keijzer, C.J., Stiekema, W.J., and Pereira, A. (1995b). Molecular characterization of the *CER1* gene of *Arabidopsis* involved in epicuticular wax biosynthesis and pollen fertility. **Plant Cell** 7, 2115-2127.
- Abel, S., Blázquez, M., Dangl, J., Deng, X.W., Graham, I., Harada, J., Jones, J. (2000). *Arabidopsis* Research 2000. **The Plant Cell** 12, 2302-2309.
- Acarkan, A., Roßberg, M., Koch, M., Schmidt, R. (2000). Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*. **The Plant Journal** 23(1), 55-62.
- Adam, R.D., Nash, T.E., Wellems, T.E. (1991). Telomeric location of *Giardia* rDNA genes. **Mol. Cell Biol.** 11(6), 3326-3330.
- Adams, G.M.W., Huang, B., and Luck, D.J.L. (1982). Temperature-sensitive assembly-defective flagella mutants of *Chlamydomonas reinhardtii*. **Genetics** 100, 579-586.
- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, C.R., Merrill, C.R., Wu, A., Olde, B., Moreno, R.F. et al., 1991. Complementary DNA sequencing: expressed sequence tags and the human genome project. **Science** 252, 1651-1656.
- Adams et al. (2000). The Genome Sequence of *Drosophila melanogaster*. **Science** 287 (5461), 2185 – 2195.
- Aharoni, A., Keizer, L.C., Bouwmeester, H.J., Sun, Z., Alvarez-Huerta, M., Verhoeven, H.A., Blaas, J., van Houwelingen, A.M., De Vos, R.C., van der Voet, H., Jansen, R.C., Guis, M., Mol, J., Davis, R.W., Schena, M., van Tunen, A.J. & O'Connell A.P. (2000). Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. **The Plant Cell** 12, 647-662.
- Aharoni, A., Vorst, O. (2001). DNA microarrays for functional plant genomics. **Plant Molec.Biol.** 48, 99-118.
- Ahmadian, A., Lundeberg, J., Nyren, P., Uhlen, M., and Ronaghi, M. 2000b. Analysis of the p53 tumor suppressor gene by pyrosequencing. **BioTechniques** 28, 140-144.
- Ahr, A., Holtrich, U., Solbach, C., Scharl, A., Strebhardt, K., Karn, T., Kaufmann, M. (2001). Molecular classification of breast cancer patients by gene expression profiling. **J. Pathol.** 195, 312-320.
- Alba, R. et al. (2004). ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. **The Plant Journal** 39,697-714.
- Alberte, R.S., Hesketh, J.D., Hofstra, G., Thornber J.P., Naylor, A.W., Bernard, R.L., Brim, C., Endrizzi, J., Kohel, R.J. (1974). Composition and activity of the photosynthetic apparatus in temperature-sensitive mutants of higher plants. **Plant Cell Physiol.** 44 (6), 588-596.
- Alderborn, A., Kristofferson, A., and Hammerling, U. 2000. Determination of single nucleotide polymorphisms by real-time pyrophosphate DNA sequencing. **Genome Res.** 10, 1249-1258.
- Alleman, M., Doctor, J. (2000). Genomic imprinting in plants: observations and evolutionary implications. **Plant Molec.Biol.** 43, 147-161.
- Allemeersch, J., Durinck, S., Vanderhaeghen, R., Alard, P., Maes, R., Seeuws, K., Bogaert, T., Coddens, K., Deschouwer, K., Van Hummelen, P., Vuylsteke, M. et al. (2005). Benchmarking the CATMA microarray. A novel tool for *Arabidopsis* transcriptome analysis. **Plant Physiology**, 137, 588-601.
- Allen, J. S., Filip, C.C., Gustafson, R.A., Allen, R.G., Walker, J.R. (1974). Regulation of bacterial cell division: genetic and phenotypic analysis of temperature-sensitive, multinucleate, filament-forming mutants of *Escherichia coli*. **J. Bacteriol.** 117, 978-986.
- Allen, K.N. and Dunaway-Mariano, D. (2004). Phosphoryl group transfer: evolution of a catalytic scaffold. **Trends in Biochem.Sci.** 29, 495-503.
- Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. **Science** 301, 653-657.
- Alonso-Blanco, C. and Koornneef, M. (2000). Naturally occurring variation in *Arabidopsis*: an underexploited resource for plant genetics. **Trends in Plant Science** 5 (1), 1360-1385.

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucl. Acids Res.** 25, 3389-3402.
- Alwine, J.C., Kemp, D.J., Stark, G.R. (1977). **Proc. Natl. Acad. Sci. U. S. A.** 74, 5350-5354.
- Akerley, B.J., Rubin, E.J., Camilli, A., Lampe, D.J., Robertson, H.M., and Mekalanos, J.J. (1998). Systematic identification of essential genes by *in vitro* mariner mutagenesis. **Proc Natl Acad Sci USA** 43, 8927-8932.
- Anderson, M.A. et al. (2001). Hydrolysis of phosphotriesters: determination of transition states in parallel reactions by heavy-atom isotope effects. **J. Am. Chem. Soc.** 123, 9246-9253.
- Andrulis, E., Neiman A.M., Zappulla D.C., and Sternglanz, R. (1998). Perinuclear localization of chromatin facilitates transcriptional silencing. **Nature** 394: 592-595.
- Antal, Z.V., Rasclé, C.V., Fèvre, M.V., Bruel, C.V. (2004). Single oligonucleotide nested PCR: a rapid method for the isolation of genes and their flanking regions from expressed sequence tags. **Current Genetics** 46 (4), 240 – 246.
- Aramli, L.A., Teschke, C.M. (1999). Single amino acid substitutions globally suppress the folding defects of temperature-sensitive folding mutants of phage P22 coat protein. **J. Biol. Chem.** 274 (32), 22217-22224.
- Aravind, L. (1998). An evolutionary classification of the metallo-beta lactamase fold proteins. **In Silico Biol.** 1, 8.
- Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W., Camilleri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J. & Williamson R.E. (1998). Molecular analysis of cellulose biosynthesis in *Arabidopsis*. **Science** 279, 717-720.
- Arkov, A.L., Freistroffer, D.V., Ehrenberg, M., Murgola, E.J. (1998). Mutations in RNAs of both ribosomal subunits cause defects in translation termination. **The EMBO Journal** 17(5), 1507-1514.
- Arondel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H.M., Somerville, C.R. (1992). Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. **Science** 258, 1353-1355.
- Ashburner, M. (1990). In: **Drosophila, A Laboratory Handbook**. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Aubourg, S., Boudet, N., Kreis, M., Lecharny, A. (2000). In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. **Plant Molec.Biol.** 42, 603-613.
- Audic, S. and J.-M. Claverie. (1997). The significance of digital gene expression profiles. **Genome Res.** 7, 986-995.
- Austin, M.J., Muskett, P., Kahn, K., Feys, B.J., Jones, J.D., and Parker, J.E. (2002). Regulatory role of SGT1 in early R gene-mediated plant defenses. **Science** 295, 2077-2080.
- Ayles, G., Sanders, T.G., Kiefer, B.J., Suzuki, D.T. (1973). Temperature-sensitive mutations in *Drosophila melanogaster*. XI. Male sterile mutants of the Y chromosome. **Devel. Biol.** 32, 239-257.
- Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K., and Schulze-Lefert, P. (2002). The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance. **Science** 295, 2073-2076.
- Aziproz-Leehan, R., Feldmann, K.A. (1997). T-DNA insertion mutagenesis in *Arabidopsis*: going back and forth. **Trends in Genetics** 13(4), 152-156.
- B**achem, C.W.B., van der Hoeven, R.S., de Bruijn, S.M., Vreugdenhil, D., Zabeau, M. & Visser, R.G.F. (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. **Plant J.** 9, 745-753.
- Bachem, C.W.B., Oomen, R.J.F.J., Visser, R.G.F. (1998). Transcript imaging with cDNA-AFLP: a step-by-step protocol. **Plant Molec.Biol.Reporter** 16, 157-173.
- Bachmair, A., Finley, D., Varshavsky, A. (1986). *In vivo* half-life of a protein is a function of its amino-terminal residue. **Science** 234, 179-186.

- Bachmair, A. and Varshavsky, A. (1989). The degradation signal in a short-lived protein. **Cell** 56 (6), 1019-1032.
- Baillie, D., Suzuki, D.T., Tarasoff, M. (1968). Temperature-sensitive mutations in *Drosophila melanogaster*. II. Frequency among second chromosome recessive lethals induced by ethyl methanesulfonate. **Can. J. Genet. Cytol.** 10: 412-420.
- Ball, C.A. et al. (2001). *Saccharomyces* genome database provides tools to survey gene expression and functional analysis data. **Nucleic Acids Research** 29(1), 80-81.
- Bancroft, I. (2000). Insights into the structural and functional evolution of plant genomes afforded by the nucleotide sequences of chromosomes 2 and 4 of *Arabidopsis thaliana*. **Yeast** 17: 1-5.
- Bancroft, I. (2001). Duplicate and diverge: the evolution of plant genome microstructure. **Trends in Genetics** 17, 89-93.
- Bancroft, I. and Dean, C. (1993a). Factors affecting the excision frequency of the maize transposable element Ds in *Arabidopsis thaliana*. **Mol. Gen. Genet.**
- Bartee, L., Malagnac, F. and Bender, J. (2001). *Arabidopsis cmt3* chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. **Genes Dev.** 15, 1753-1758.
- Baskin, T.I. (2001). On the alignment of cellulose microfibrils by cortical microtubules: A review and a model. **Protoplasma** 215, 150-171.
- Bassett Jr., D.E., Eisen, M.B., Boguski, M.S. (1999). Gene expression informatics - it's all in your mine. **Nature Genetics Suppl.** 21, 51-55.
- Baulcombe, D.C. (1996) RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. **Plant Mol. Biol.** 32, 79-88.
- Beach, D., Durkacz, B. & Nurse, P. (1982). Functionally homologous cell cycle control genes in budding and fission yeast. **Nature** 30, 706-709.
- Beadle, G.W. and Tatum, E.L. (1945). Neurospora. II. Methods of producing and detecting mutants concerned with nutritional requirements. **Am. J. Botany** 32, 678-686.
- Beemster, G.T.S., Baskin, T.I. (1998). Analysis of cell division and elongation underlying the developmental acceleration of root growth in *Arabidopsis thaliana*. **Plant Physiol.** 116, 1515-1526.
- Behm, M., Lowman, H., Ng, S.C., Bina, M. (1988). Analysis of temperature-sensitive mutations in the simian virus 40 gene encoding virion protein 1. **Proc. Natl. Acad. Sci. USA** 85, 9421-9425.
- Bell, C.J. and Ecker, J.R. (1994). Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. **Genomics** 19 (1), 137-144.
- Benard, C., McCright, B., Zhang, Y., Felkai, S., Lakowski, B. & Hekimi S. (2001). The *C. elegans* maternal-effect gene *clk-2* is essential for embryonic development, encodes a protein homologous to yeast Tel2p and affects telomere length. **Development** 128, 4045-4055.
- Benfey, P.N., Linstead, P.J., Roberts, K., Schiefelbein, J.W., Hauser, M.T., Aeschbacher, R.A. (1993). Root development in *Arabidopsis*: four mutants with dramatically altered root morphogenesis. **Development** 119, 57-70.
- Berna, G., Robles, P., Micol, J.L. (1999) A mutational analysis of leaf morphogenesis in *Arabidopsis thaliana*. **Genetics** 152, 729-742.
- Bernstein, H., Edgar, R.S., Denhardt, G.H. (1965). Intragenic complementation among temperature-sensitive mutants of bacteriophage T4D. **Genetics** 51, 987-1002.
- Bertioli, D.J. et al. (1995). An analysis of differential display shows a strong bias towards high copy number mRNAs. **Nucleic Acids Res.** 23, 4520-4523.
- Bevan, M. et al. (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. **Nature** 391, 485-488.

- Bevan, M., Bancroft, I., Mewes, H.W., Martienssen, R. & McCombie, R. (1999). Clearing a path through the jungle: progress in *Arabidopsis* genomics. **Bioessays** 21, 110-120.
- Bevins, M.A., Madhavan, S., Markwell, J. (1993) Two sweetclover (*Melilotus alba* Desr.) mutants temperature-sensitive for chlorophyll expression. **Plant Physiology** 103(4), 1123-1131.
- Bewley, J.D., Burton, R.A., Morohashi, Y., Fincher, G.B. (1997). Molecular cloning of a cDNA encoding a (1→4) β -mannan endohydrolase from the seeds of germinated tomato (*Lycopersicon esculentum*). **Planta** 203, 454-459.
- Blackburn, E.H., (1991). Structure and function of telomeres. **Nature** 350, 569-573.
- Bleecker, A.B., Estelle, M.A., Somerville, C. and Kende, H. (1988). Insensitivity to ethylene conferred by a dominant mutation in *Arabidopsis thaliana*. **Science** 241, 1086-1089.
- Bohnert, H.J., Nelson, D.E., Jensen, R.G. (1995). Adaptations to environmental stresses. **Plant Cell**, 7, 1099-1111.
- Bonaldo, M.F., Lennon, G., and Soares, M.B. (1996). Normalization and subtraction: Two approaches to facilitate gene discovery. **Genome Res.** 6, 791-806.
- Bonin, C.P., Potter, I., Vanzin, G.F., Reiter, W-D. (1997). The *MUR1* gene of *Arabidopsis thaliana* encodes an isoform of GDP-D-mannose-4,6-dehydratase, catalyzing the first step in the *de novo* synthesis of GDP-L-fucose. **Proc. Natl. Acad. Sci. USA** 94, 2085-2090.
- Borevitz, J.O., Chory, J. (2004). Genomics tools for QTL analysis and gene discovery. **Curr.Opin.Plant Biol.** 7,132-136.
- Borevitz, J.O., Liang, D., Plouffe, D., Chang, H.S., Zhu, T., Weigel, D., Berry, C.C., Winzeler, E., Chory J. (2003). Large-scale identification of single-feature polymorphisms in complex genomes. **Genome Research** 13, 513-523.
- Botstein D., White R.L., Skolnick M., Davis R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **Am. J. Hum.Genet.** 32, 314-331.
- Bouché, N. and Bouchez, D. (2001). *Arabidopsis* gene knockout: phenotypes wanted. **Curr. Opin. Plant Biol.** 4, 111-117.
- Bouchez, D. & Höfte H. (1998). Functional genomics in plants. **Plant Physiol.** 118, 725-732.
- Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. (1989). Genes directing flower development in *Arabidopsis*. **The Plant Cell** 1, 37-52.
- Bowtell, D.D.L. (1999). Options available - from start to finish - for obtaining expression data by microarray. **Nature Genetics Suppl.** 21, 25-32.
- Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., Görlach, J. (2001). Growth stage-based Phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. **The Plant Cell** 13,1499-1510.
- Bram, R.J. and Kornberg, R.D. (1985). Specific protein binding to far upstream activating sequences in polymerase II promoters. **Proc. Natl. Acad. Sci. USA** 82 (1), 43-47.
- Brand, A. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. **Development** 118, 401-415.
- Brenner, S., Williams, S.R., Vermaas, E.H., Storck, T., Moon, K., McCollum, C., Mao, J.I., Luo, S., Kirchner, J.J., Eletr, S., et al. (2000a). *In vitro* cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. **Proc. Natl. Acad. Sci. USA** 97, 1665-1670.
- Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., et al. (2000b). Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. **Nat. Biotechnol.** 18, 630-634.
- Breton, A.H. and Sung, Z.R. (1982). Temperature-sensitive carrot variants impaired in somatic embryogenesis. **Dev. Biol.** 90, 58-66.

- Breyne, P. and Zabeau, M. (2001). Genome-wide expression analysis of plant cell cycle modulated genes. **Curr. Opin. Plant Biol.** 4, 136-142.
- Brown Jr, R.M. (1999). Cellulose structure and biosynthesis. **Pure Appl.Chem.** 71, 767-775.
- Brown D.D., Dawid I.B. (1969). Developmental Genetics. **Ann. Rev. Genet.** 3, 127.
- Browse, J., McCourt, P. and Somerville, C.R. (1985). A mutant of *Arabidopsis* lacking a chloroplast-specific lipid. **Science** 227, 763-765.
- Bruce, W., Folkerts, O., Garnaat, C., Crasta, O., Roth, B., Bowen, B. (2000). Expression profiling of the maize flavonoid pathway genes controlled by estradiol-inducible transcription factors CRC and P. **Plant Cell** 12, 65-79.
- Burn, J.E., Hocart, C.H., Birch, R.J., Cork, A.C., Williamson, R.E. (2002). Functional analysis of the cellulose synthase genes *CesA1*, *CesA2*, and *CesA3* in *Arabidopsis*. **Plant Physiology** 129, 797-807.
- Burn, J.E., Hurley, U.A., Birch, R.J., Arioli, T., Cork, A., Williamson, R.E. (2002). The cellulose deficient *Arabidopsis* mutant *rsw3* is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. **Plant J.** 32, 949-960.
- C**amilleri, C., Lafleur, J., Macadré, C., Varoquaux, F., Parmentier, Y., Picard, G., Caboche, M. and Bouchez, D. (1998). A YAC contig map of *Arabidopsis thaliana* chromosome 3. **The Plant Journal** 14(5), 633 - 642.
- Campos-De Quiroz, H. (2002). Plant genomics: an overview. **Biol Res** 35, 385-399.
- Carrels, N., Bernardi, G. (2000). Two classes of Genes in Plants. **Genetics** 154, 1829-1825.
- Carl, P. L. (1970). *Escherichia coli* mutants with temperature-sensitive synthesis of DNA. **Mol. Gen. Genet.** 109, 107-122.
- Carulli, J.P. et al. (1998). High throughput analysis of differential gene expression. **Journ. of Cell.Biochem.** 72,S30-31, 286-296.
- Cashmore, A.R. (2003). Cryptochromes: enabling plants and animals to determine circadian time. **Cell** 114, 537-543.
- Casimiro, I., Beeckman, T., Graham, N., Bhalerao, R., Zhang, H., Casero, P., Sandberg, G., Bennett, M.J. (2003). Dissecting *Arabidopsis* lateral root development. **Trends in Plant Science** 8,165-171.
- Cassada, R., Isnenghi, E., Culotti, M., von Ehrenstein, G. (1981). Genetic analysis of temperature-sensitive embryogenesis mutants in *Caenorhabditis elegans*. **Dev. Biol.** 84, 193-205.
- Catrina, I.E. and Hengge, A.C. (2003). Comparisons of phosphorothioate with phosphate transfer reactions for a monoester, diester, and triester: isotope effect studies. **J. Am. Chem. Soc.** 125, 7546-7552.
- Causse, M.A., Fulton, T.M., Cho, Y.G., Ahn, S.N., Chungwongse, J. et al. (1994). Saturated molecular map of the rice genome based on an interspecific backcross population. **Genetics** 138,1251-1274.
- Cayrol, C., Cabrol, G., Ducommun, B. (1997) Use of the two-hybrid system to identify protein-protein interaction temperature-sensitive mutants: application to the CDK2/p21Cip1 interaction. **Nucleic Acids Res** 25, 3743-3744.
- Celenza, J.L., Jr., Grisafi, P.L., Fink, G.R. (1995). A pathway for lateral root formation in *Arabidopsis thaliana*. **Genes Dev.** 9, 2131-2142.
- Celis, J.E., Kruhoffer, M., Gromova, I., Frederiksen, C., Østergaard, M., Thykjaer, T., Gromov, P., Yu, J., Palsdottir, H., Magnusson, N., Ornótt, T.F., 2000. Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics. **FEBS Lett.** 480 (1), 2-16.
- Cha, R.S., Zarbl, H., Keohavong, P., Thilly, W.G. (1992). Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. **PCR Methods Appl.** 2, 14-20.
- Chakshumathi, G., Mondal, K., Lakshmi, G.S., Singh, G., Roy, A., babu CH., R., Madhusudhanan, S., Varadarajan, R. (2004). Design of temperature-sensitive mutants solely from amino acid sequence. **Proc. Natl. Acad. Sci. USA** 101(21), 7925-7930.

- Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X.C., Stern, D., Winkler, J., Lockhart, D.J., Morris, M.S., Fodor, S.P.A. (1996). Accessing genetic information with high-density DNA arrays. **Science** 274, 610-613.
- Chen, X., Livak, K.J., Kwok, P.Y. (1998). A homogeneous, ligase-mediated DNA diagnostic test. **Genome Res.** 8, 549-546.
- Cheong, Y.H., Chang, H.S., Gupta, R., Wang, X., Zhu, T., Luan, S. (2002). Transcriptional profiling reveals novel interactions between, wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. **Plant Physiology** 129, 661-667.
- Cho, R.J. et al. (1998). A genome-wide Transcriptional analysis of the mitotic cell cycle. **Molecular Cell** 2, 65-73.
- Cho, R.J., Mindrinos, M., Richards, D.R., Sapolsky, R.J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T.L., Stammers, M., Federspiel, N., et al. (1999). Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. **Nat. Genet.** 23: 203-207.
- Chong, S., Williams, K.S., Wotkowicz, C., Xu, M-Q. (1998). Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar ATPase (VMA) intein. **J. Biol. Chem.** 273, 10567-10577.
- Clark, S.E., Jacobsen, S.E., Levin, J.Z., Meyerowitz, E.M. (1996). The *CLAVATA* and *SHOOTMERISTEMLESS* loci competitively regulate meristem activity in *Arabidopsis*. **Development** 122, 1567-1575.
- Clark, M., Panopoulou, G., Cahill, D., Büssov, K., Lehrach, H. (1999). Construction and analysis of arrayed cDNA libraries. **Methods Enzymol.** 303, 205-233.
- Chuang, C.F. and Meyerowitz, E.M. (2000). Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. **Proc. Natl. Acad. Sci. USA** 97 (9), 4985-4990.
- Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. **Plant J.** 16, 735-743.
- Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L., Henikoff S. (2001). High-throughput screening for induced point mutations. **Plant Physiol.** 126, 480-484.
- Cnops, G., Wang, X., Linstead, P., Van Montagu, M., Van Ijsebettens, M., Dolan, L. (2000). *TORNADO1* and *TORNADO2* are required for the specification of radial and circumferential pattern in the *Arabidopsis* root. **Development** 127, 3385-3394.
- Cnops, G., Jover-Gil, S., Peters, J.L., Neyt, P., De Block, S., Robles, P., Ponce, M.R., Gerats, T., Micol, J.L., Van Ijsebettens, M. (2004) The *rotunda2* mutants identify a role for the *LEUNIG* gene in vegetative leaf morphogenesis. **Journal of Experimental Botany** 402, 1529-1539.
- Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. **Nature** 353, 31-37.
- Collet, J.F., Stroobant, V., Pirard, M., Delpierre, G., Van Schaffingen, E. (1998). A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX (T/V) motif. **The Journal of Biol.Chem.** 273(23), 14107-14112.
- Conklin, P.L., Williams, E.H., Last, R.L. (1996). Environmental stress sensitivity of an ascorbic acid-deficient *Arabidopsis* mutant. **Plant Biology** 93, 9970-9974.
- Conklin, P.L., Pallanca, J.E., Last, R.L., Smirnov, N. (1997) L-Ascorbic acid metabolism in the ascorbate-deficient *Arabidopsis* mutant *vtc1*. **Plant Physiol.** 115, 1277-1285.
- Conklin, P.L., Norris, S.R., Wheeler, G.L., Williams, E.H., Smirnov, N., Last, R.L. (1999). Genetic evidence for the role of GDP-mannose in plant ascorbic acid (vitamin C) biosynthesis. **Plant Biology** 96, 4198-4203.
- Conklin, P.L., Saracco, S.A., Norris, S.R., Last, R.L. (2000). Identification of ascorbic acid-deficient *Arabidopsis thaliana* mutants. **Genetics** 154, 847-856.
- Cooke, R. et al. (1996). Further progress towards a catalogue of all *Arabidopsis* genes: analysis of a set of 5000 non-redundant ESTs. **The Plant Journal** 9(1), 101-124.

- Copenhaver, G.P., Doelling, J.H., Gens, J.S., Pikaard, C.S. (1995). Use of RFLPs larger than 100 kbp to map the position and internal organization of the nucleolar organizer region on chromosome 2 in *Arabidopsis thaliana*. **Plant J.** 7(2), 273-286.
- Copenhaver, G.P., Pikaard, C.S. (1996). RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolar organizer regions of *Arabidopsis thaliana* adjoin the telomeres on chromosomes 2 and 4. **Plant J.** 9(2), 259-272.
- Couee, I., Sulmon, C., Gouesbet, G., El Amrani, A. (2006). Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. **J. Exp. Bot.** 57(3), 449-459.
- Coughlan, S.J., Agrawal, V., Meyers, B. (2004). A comparison of global gene expression measurement technologies in *Arabidopsis thaliana*. **Comp. Funct. Genom.** 5, 245-252.
- Cox, J. H. and Strack, H.B. (1971). Cold-sensitive mutants of bacteriophage Lambda. **Genetics** 67, 5-17.
- Crick, F. H. C., and Orgel, L.E. (1964). The theory of interallelic complementation. **J. Mol. Biol.** 8, 161-165.
- Crowe, M.L. et al. (2003). CATMA: a complete *Arabidopsis* GST database. **Nucleic Acids Res.** 31, 156-158.
- D**atson, N.A., van der Perk-de Jong, J., van den Berg, M.P., de Kloet, E.R., Vreugdenhil, E. (1999). MicroSAGE: A modified procedure for serial analysis of gene expression in limited amounts of tissue. **Nucleic Acids Res.** 27, 1300-1307.
- Dean, C., Schmidt, R. (1995). PLANT GENOMES : A Current Molecular Description. **Annu.Rev.Plant Physiol. Plant Mol.Biol.** 46, 395-418.
- Dear, P.H. 1997. HAPPY mapping. In: **Genome mapping: A practical approach**, 95-124; edited by P. H. Dear. IRL Press, Oxford, UK.
- De Buck, S., De Wilde, C., van Montagu, M., Depicker, A. (2000). T-DNA vector backbone sequences are frequently integrated into the genome of transgenic plants obtained by *Agrobacterium*-mediated transformation. **Mol. Breed.** 6, 459-468.
- De Castro, R.D., van Lammeren, A.A.M., Groot, S.P.C., Bino, R.J., Hilhorst, H.W.M. (2000). Cell division and subsequent radicle protrusion in tomato seeds are inhibited by osmotic stress but DNA synthesis and formation of microtubular cytoskeleton are not. **Plant Physiology** 122, 327-335.
- Dellaqi, A., Birch, P., Heilbronn, J., Lyon, G., Toth, I. (2000). cDNA-AFLP analysis of differential gene expression in the prokaryotic plant pathogen *Erwinia carotovora*. **Microbiology**, 146, 165-171.
- del Pozo, J.C., Allona, I., Rubio, V., Leyva, A., de la Peña, A., Aragoncillo, C., Paz-Ares, J. (1999). A type 5 acid phosphatase gene from *Arabidopsis thaliana* is induced by phosphate starvation and by some other types of phosphate mobilising/oxidative stress conditions. **Plant Journal** 19, 579-589.
- De Paepe, A., Vuylsteke, M., Van Hummelen, P., Zabeau, M. Van Der Straeten, D. (2004). Transcriptional profiling by cDNA-AFLP and microarray analysis reveals novel insights into the early response to ethylene in *Arabidopsis*. **The Plant Journal** 39,537-559.
- DeRisi, J.L., Iyer, V.R. & Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. **Science** 278, 680-686.
- De Sanctis, V., Bertozzi, C., Costanzo, G., Di Mauro, E., and Negri, R. (2001). Cell cycle arrest determines the intensity of the global transcriptional response of *Saccharomyces cerevisiae* to radiation. **Radiat. Res.** 156, 379-387.
- Desikan, R., Mackerness, S.A.H., Hancock, J.T., Neill, S.J. (2001). Regulation of the *Arabidopsis* transcriptome by oxidative stress. **Plant Physiology** 127, 159-172.
- De Smet, F., Mathys, J., Marchal, K., Thijs, G., De Moor, B. & Moreau Y. (2002). Adaptive quality-based clustering of gene expression profiles. **Bioinformatics** 18, 735-746.
- Desprez, T., Amselem, J., Caboche, M. & Höfte, H. (1998). Differential gene expression in *Arabidopsis* monitored using cDNA arrays. **Plant J.** 14(5), 643-652.

- Dewaard, P.A., Paul, A.V., Lehman, I.R. (1965). The structural gene for deoxyribonucleic acid polymerase in bacteriophage T4 and T5. **Proc. Natl. Acad. Sci. USA** 54, 1241-1248.
- D'Haeseleer, P., Liang, S., Somogyi, R. (2000). Genetic network inference: from co-expression clustering to reverse engineering. **Bioinformatics** 16(8), 707-726.
- Diatchenko, L., Lau, Y. F. C., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D., et al. (1996). Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries **Proc. Natl. Acad. Sci. USA** 93, 6025-6030.
- Drmanac, R. and Drmanac, S. (1999). cDNA screening by array hybridization. **Methods Enzymol.** 303, 165-178.
- Dohmen, R.J., Wu, P., Varshavsky, A. (1994). Heat-inducible degtron: a method for constructing temperature-sensitive mutants. **Science** 263, 1273-1276.
- Donson, J. et al. (2002). Comprehensive gene expression analysis by transcript profiling. **Plant Molec. Biol.** 48, 75-97.
- Drenkard, E., Richter, B.G., Rozen, S., Stutius, L.M., Angell, N.A., Mindrinos, M., Cho, R.J., Oefner, P.J., Davis, R.W., Ausubel, F.M. (2000). A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in *Arabidopsis*. **Plant Physiol.** 124, 1483-1492.
- Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P., Trent, J.M. (1999). Expression profiling using cDNA microarrays. **Nature Genetics Suppl.** 21, 10-14.
- Dunwell, J.M., Moya-León, M.A., Herrera, R. (2001) Transcriptome analysis and crop improvement. **Biol. Res.** 34 3-4, 153-164.
- Durrant, W.E., Rowland, O., Piedras, P., Hammond-Kosack, K.E., Jones, J.D. (2000). cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. **The Plant Cell** 12(6), 963-977.
- Dzeja, P.P. and Terzic, A. (2003). Phosphotransfer networks and cellular energetics. **J. Exp. Biol.** 206, 2039-2047.

Edgar, R. S. and Lielausis, I. (1964). Temperature-sensitive mutants of bacteriophage T4D: their isolation and genetic characterization. **Genetics** 49, 649-62.

- Eisen, M. and Brown, P. (1999) DNA arrays for analysis of gene expression. **Methods Enzymol.** 303, 179-205.
- Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. **Proc. Natl. Acad. Sci. U S A** 95, 14863-14868.
- Eisenberg, D., Schwarz, E., Komaromy, M., Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. **J. Mol. Biol.** 179 (1), 125-142.
- Eleftheriou, E.P., Baskin, T.I., Hepler, P.K. (2005). Aberrant cell plate formation in the *Arabidopsis thaliana* microtubule organization 1 mutant. **Plant and Cell Physiology** 46(4), 671-675.
- Emons, A.M.C., Derksen, J. and Sassen, M.M.A. (1992). Do microtubules orient plant cell wall microfibrils? **Physiol. Plant.** 84, 486-493.
- Epstein, R., Bolle H., Steinberg C.M., Kellenberger E., Boy de la Tour E., Chevalley R., Edgar R.S., Susman M., Denhardt G.H., Lielausis H. (1963) Physiological studies of conditional lethal mutants of bacteriophage T4D. **Cold Spring Harbor Symp. Quant. Biol.** 28, 375-392.
- Ermolaeva, O. et al. (1998). Data management and analysis for gene expression arrays. **Nature genetics** 20, 19-23.
- Errampalli, D., Patton, D., Castle, L., Mickelson, L., Hansen, K., Schnall, J., Feldmann, K., Meinke, D. (1991). Embryonic lethals and T-DNA insertional mutagenesis in *Arabidopsis*. **Plant Cell** 3 (2) 149-157.
- Esposito, M. S., and Esposito, R.E. (1969). The genetic control of sporulation in *Saccharomyces*. I. The isolation of temperature-sensitive sporulation-deficient mutants. **Genetics** 61, 79-89.

Ewing, R., O. Piroët, and J.-M. Claverie. (1999). Comparative analysis of the *Arabidopsis* and rice expressed sequence tag (EST) sets. **Silico Biol.** 1, 18.

Fane, B. and King, J. (1991) Intragenic suppressors of folding defects in the P22 tailspike protein. **Genetics** 127, 263-277.

Fane, B., Villafane, R., Mitraki, A., King, J. (1991). Identification of global suppressors of temperature-sensitive folding mutants of the P22 tailspike protein. **J. Biol. Chem.** 262, 11640-11648.

Fangman, W.L. and Novick, A. (1968). Characterization of two bacterial mutants with temperature-sensitive synthesis of DNA. **Genetics** 60, 1-17.

Fattig, W.D., and Rickoll, W.L. (1972). Isolation of recessive third chromosome temperature-sensitive mutants in *Drosophila melanogaster*. **Genetics** 71, 309-313.

Fedor, M.J., Lue, N.F., Kornberg, R.D. 1988. Statistical positioning of nucleosomes by specific protein-binding to an upstream activating sequence in yeast. **J. Mol. Biol.** 204: 109-127.

Fedoroff, N.V. (1989). About maize transposable elements and development. **Cell** 56 (2), 181-91.

Fedoroff N.V., Smith D.L. (1993) A versatile system for detecting transposition in *Arabidopsis*. **Plant J** 3, 273-289.

Feldmann, K.A. (1992). T-DNA insertion mutagenesis in *Arabidopsis*: Seed infection/transformation. In: **Methods in Arabidopsis Research**, 274-289. Edited by C. Koncz, N.-H. Chua, and J. Schell. Singapore: World Scientific.

Feldmann K.A., Malmberg R.L., Dean C. (1994). Mutagenesis in *Arabidopsis*. In: **Arabidopsis**, 137. Cold Spring Harbor Laboratory Press.

Fincham, J.R.S. (1966). **Genetic Complementation**. Benjamin, New York.

Fletcher, J.C. (2002). Shoot and Floral Meristem Maintenance in *Arabidopsis*. **Annu.Rev.Plant.Biol.** 53,45-66.

Fodor, S., Read, J., Pirrung, M., Stryer, L., Lu, A., Solas, D. (1991). Light-directed, spatially addressable parallel chemical synthesis. **Science**, 251, 767-773.

Fodor, S., Rava, R., Huang, X., Pease, A., Holmes, C., Adams, C. (1993). Multiplexed biochemical assays with biological chips. **Science**, 364, 555-556.

Foster, G. and Suzuki, D.T. (1970). Temperature-sensitive mutations in *Drosophila melanogaster*, IV. A mutation affecting eye facet arrangement in a polarized manner. **Proc. Natl.Acad. Sci. USA** 67, 738-745.

Fracheboud, Y., King, P.J. (1991). An auxin-auxotrophic mutant of *Nicotiana plumbaginifolia*. **Mol. Gen. Genet.** 227, 397-400.

Fransz, P., Armstrong, S., Alonso-Blanco, C., Fischer, T.C., Torres-Ruiz, R.A., Jones, G. (1998). Cytogenetics for the model system *Arabidopsis thaliana*. **The Plant Journal** 13(6), 867-876.

Frey, P.A. et al. (1987). Pathways of enzymatic phosphotransfer reactions. **Phosphorus Sulfur Silicon Related Elements** 30, 563-566.

Fromont-Racine, M., Rain, J.C., Legrain, P. (1997). Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. **Nature Genetics** 16, 277-282.

Fukaki, H., Wysocka-Diller, J., Kato, T., Fujisawa, H., Benfey, P.N., Tasaka, M. (1998). Genetic evidence that the endodermis is essential for shoot gravitropism in *Arabidopsis thaliana*. **Plant J.** 14: 425-430.

Galitski, T., Saldanha, A.J., Styles, C.A., Lander, E.S. and Fink, G.R. (1999). Ploidy regulation of gene expression. **Science** 285, 251-254.

Ge, S. J., Yao, X. L., Yang, Z. X. & Zhu, Z. P. (1998). An *Arabidopsis* embryonic lethal mutant with reduced expression of alanyl-tRNA synthetase gene. **Cell Res.** 8, 119-134.

- Gelbart, W.M. (1998). Databases in genomic research. **Science** 282, 659-661.
- Ghosh, D. (2000). High throughput and global approaches to gene expression. **Comb.Chem. & High Throughput Screening** 3, 411-420.
- Giaever, G., Chu, A.M., Ni, L., Connelly, C. et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. **Nature** 418, 387-391.
- Gibson, G., Dworkin, I. (2004). Uncovering cryptic genetic variation. **Nature reviews/Genetics** 5, 681-690.
- Gichner, T. and Veleminsky, J. (1977). **Biologia Plantarum (Praha)** 19, 153.
- Giddings, T.H. and Staehelin, L.A. (1991). Microtubule-mediated control of microfibril deposition; a re-examination of the hypothesis. In: **The Cytoskeletal Basis of Plant Growth and Form** (Lloyd, C.W., ed.). San Diego, CA: Academic Press, 85-99.
- Golub, T.R., Slonim, D.K., Tamayo, P., Huard, P., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. **Science** 286 (5439), pp. 531-537.
- Gosti, F., Bertauche, N., Vartanian, N., and Giraudat, J. (1995). Absciscic acid dependent and independent regulation of gene expression by progressive drought in *Arabidopsis thaliana*. **Mol. Gen. Genet.** 246, 10-18.
- Gilchrist, E.J., Haughn, G.W. (2005). TILLING without a plough: a new method with applications for reverse genetics. **Curr.Opin.Plant Biol.** 8, 1-5.
- Gillespie D, Spiegelman S. (1965). A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. **J. Mol. Biol.** 12 (3), 829-842.
- Gillmor, C.S., Poindexter, P., Lorieau, J., Palcic, M.M., Somerville, C. (2002). α -glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*. **The Journal of Cell Biol.** 156, 1003-1013.
- Giniger, E., Varnum, S. M., Ptashne, M. (1985). Specific DNA binding of GAL4, a positive regulatory protein of yeast. **Cell** 40, 767-774.
- Giovanonni, J., Wing, R. and Tanksley, S.D. (1992). Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. **Nucl. Acid Res.** 19, 6553-6558.
- Giraudat, J., Hauge, B.M., Valon, C., Smalle, J., Parcy, F., Goodman, H.M. (1992). Isolation of the *Arabidopsis* *ABI3* gene by positional cloning. **Plant Cell** 4, 1251-1261.
- Giuliano, G., Lo Schiavo, F., Terzi, M. (1984). Isolation and developmental characterization of temperature-sensitive carrot cell variants. **Theor. Appl. Genet.** 67, 179-183.
- Glazebrook, J., Rogers, E.E. and Aushubel, F.M. (1997). Use of *Arabidopsis* for genetic dissection of plant defense responses. **Annu. Rev. Genet.** 31, 547-569.
- Gonda, D.K., Bachmair, A., Wuenning, I., Tobias, J.W., Lane, W.S., Varshavsky, A. (1989). Universality and structure of the N-end rule. **J. Biol. Chem.** 264 (28), 16700-16712.
- Goodman, H.M., Eckers, J.R., Dean C. (1995). The genome of *Arabidopsis thaliana*. **Proc.Natl.Acad.Sci.USA** 92, 10831-10835.
- Goldenberg, D. and King, J. (1981). Temperature sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein: Mature ts proteins synthesized at 30°C. **J. Mol. Biol.** 145
- Goldenberg, D.P., Smith, D.H., King, J. (1983). Genetic analysis of the folding pathway for the tail spike protein of phage P22 **Proc.Natl.Acad.Sci.USA** 80,7060-7064.
- Goldenberg, D.P. (1988). Genetic studies of protein stability and mechanisms of folding. **Annual Review of Biophysics and Biophysical Chemistry** 17, 481-507.
- Goossens, A., Häkkinen, S.T., Laakso, I., Seppänen-Laakso, T., Biondi, S., De Sutter, V., Lammertyn, F., Nuutila, A. M., Söderlund, H., Zabeau, M., et al. (2003). **Proc. Natl. Acad. Sci. USA** 100, 8595-8600.

- Gordon, C. L., and King, J. (1993). Temperature-sensitive mutations in phage P22 coat protein which interfere with polypeptide chain folding. **J. Biol. Chem.** 268, 9358-9368.
- Gordon, C.L. and King J. (1994). Genetic properties of temperature-sensitive folding mutants of the coat protein of phage P22. **Genetics** 136, 427-438.
- Gossen, M. and Bujard H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. **Proc. Natl. Acad. Sci. USA** 89 (12), 5547-51.
- Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. **Science** 268 (5218), 1766-1769.
- Goto, K. and Meyerowitz, E.M. (1994). Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. **Genes and Development** 8, 1548-1560.
- Grey, N.S., Wodicka, L., Thunnissen, A.M., Norman, T.C., Kwon, S., Espinoza, F.H., Morgan, D.O., Barnes, G., LeClerc, S., Meijer, L., Kim, S.H., Lockhart, D.J. and Schultz, P.G. (1998). Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. **Science** 281, 533-538.
- Green, E.D. (2001). Strategies for the systematic sequencing of complex genomes. **Nature Reviews Genetics** 2, 573-583.
- Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., et al. (2003). Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. **Genetics** 164, 731-740.
- Gregory, S.T., Carr, J.F., Rodriguez-Correa, D., Dahlberg, A.E. (2005). Mutational analysis of 16S and 23S rRNA genes of *Thermus thermophilus*. **Journ.of Bacteriology** 187(14), 4804-4812.
- Gresham, D., Ruderfer, D., M., Pratt, S.C., Schacherer, J., Dunham, M.J., Botstein, D., Kruglyak, L. (2006). Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray. **Science** 311 (5769), 1932 – 1936.
- Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C. & Gelbart, W.M. (1996). DNA function. In: An introduction to genetic analysis, Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C. & Gelbart, W.M. (Eds). W.H. Freeman & Co., New York, 383-420.
- Grzyska, P.K. et al. (2003). Transition state differences in hydrolysis reactions of alkyl versus aryl phosphate monoester monoanions. **J. Am. Chem. Soc.** 125, 13106–13111.
- Guthrie, C., Nashimoto, H. & Nomura, M. (1969). Structure and function of *E. coli* ribosomes. 8. Cold-sensitive mutants defective in ribosome assembly. **Proc. Natl. Acad. Sci. USA** 63, 384-391.
- Guzman, P. and Ecker, J.R. (1990). Exploiting the triple response of *Arabidopsis* to identify ethylene-related mutants. **Plant Cell** 2, 513-523.
- Gygi et al. (1999). Correlation between protein and mRNA abundance in yeast. **Mol. Cell Biol.** 19, 1720-1730.

- H**aag, E. and Raman V. (1994). Effects of primer choice and source of Taq DNA polymerase on the banding patterns of differential display RT-PCR. **Biotechniques** 17, 226–228.
- Hall, J.C. (1978a). Behavioral analysis in *Drosophila* mosaics. In: **Genetic Mosaics and Cell Differentiation**, 259-305. Edited by W. J. Gehring. Springer-Verlag, Berlin.
- Hall J.C. and Greenspan, R.J. (1979). Genetic analysis of *Drosophila* neurobiology. **Ann. Rev. Genet.** 13: 127-195.
- Harmer, S.L., Kay, S.A. (2000). Microarrays : determining the balance of cellular transcription. **The Plant Cell** 12,613-615.
- Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.-S., Han, B., Zhu, T., Wang, X., Kreps, J.A., Kay, S.A. (2000). Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. **Science** 290, 2110-2113.

- Harris, S.D. and Pringle, J.R. (1991). Genetic analysis of *Saccharomyces cerevisiae* chromosome I: on the role of mutagen specificity in delimiting the set of genes identifiable using temperaturesensitive-lethal mutations. **Genetics** 127, 279-285.
- Hartwell, L.H. (1971a). Genetic control of the cell division cycle in yeast. II. Genes controlling DNA replication and its initiation. **J. Mol. Biol.** 59: 183-194.
- Hartwell, L.H. (1967). Macromolecule synthesis in temperature-sensitive mutants of yeast. **J. Bacteriol.** 93, 1662-1670.
- Hartwell, L.H., Culotti, J. & Reid, B.J. (1970). Genetic control of the cell-division cycle in yeast. I. Detection of mutants. **Proc. Natl. Acad. Sci. USA** 66, 352-359.
- Hartwell, L.H., Culotti, J., Pringle, J.R. & Reid B.J. (1974). Genetic control of the cell division cycle in yeast. **Science** 183, 46-51.
- Haughn, G.W. and Sommerville, C. (1986). Sulfonylurea-resistant mutants of *Arabidopsis thaliana*. **Mol. Gen. Genet.** 204, 430-434.
- Hazen, S.P., Borevitz, J.O., Harmon, F.G., Pruneda-Paz, J.L., Schultz, T.F., Yanovsky, M.J., Liljegren, S.J., Ecker, J.R., Kay, S.A. (2005). Rapid Array Mapping of Circadian Clock and Developmental Mutations in *Arabidopsis*. **Plant Physiol.** 138, 990-997.
- Hecht, M. H., Sturtevant, J. M., and Sauer, R. T. (1984). Effect of Single Amino Acid Replacements on the Thermal Stability of the NH₂-terminal Domain of Phage λ Repressor. **Proc. Natl. Acad. Sci. USA** 81 (18), 5685-5689.
- Heiskanen, M., Kallioniemi, O.-P., Palotie, A. (1996). Fiber-FISH: Experiences and a refined protocol. **Genet. Anal. Biomol. Eng.** 12, 179-184.
- Heller, R.A., Schena, M., Chai, A., Shalon, D., Bedilion, T., Gilmore, J., Woolley, D.E., Davis, R.W. (1997). Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. **Proc. Natl. Acad. Sci. USA** 94, 2150-2155.
- Henikoff, S., Comai, L. (2003). Single-nucleotide mutations for plant functional genomics. **Annu.Rev.Plant.Biol.** 54, 375-401.
- Heremans, B. and Jacobs, M. (1997). A mutant of *Arabidopsis thaliana* (L.) Heynh. with modified control of aspartate kinase by threonine. **Biochem. Genet.** 35, 139-153.
- Herman, G.E. (1998). Physical mapping of the mouse genome. **Methods** 14, 135-151.
- Heyer, L.J., Kruglyak, S., Yooseph, S. (1999). Exploring expression data: identification and analysis of coexpressed genes. **Genome Research** 9,1106-1115.
- Hieter, P., Boguski, M. (1997). Functional Genomics: it's all how you read it. **Science mag.** 278, 601-602.
- Hill, J. P. and Lord, E. M. (1989). Floral development in *Arabidopsis thaliana*: a comparison of the wild type and the homeotic pistillata mutant. **Can. J. Bot.** 67, 2922-2936.
- Hilson, P., Allemeersch, J., Altmann, T., Aubourg, S., Avon, A., Beynon, J., Bhalerao, R.P., Bitton, F., Caboche, M., Cannoot, B., et al. (2004). Versatile gene-specific sequence tags for *Arabidopsis* functional genomics: transcript profiling and reverse genetics applications. **Genome Res.** 14, 2176-2189.
- Himmelsbach, R., Williamson, R.E., Wasteneys, G.O. (2003). Cellulose microfibril alignment recovers from DCB-induced disruption despite microtubule disorganization. **The Plant Journal** 36, 565-575.
- Hirsh, D. and Vanderslice, R. (1976). Temperature-sensitive developmental mutants of *Caenorhabditis elegans*. **Developmental Biology** 49 (1), 220-235.
- Höfte, H., Desprez, T., Amselem, J., Chiapello, H., Caboche, M. (1993) An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. **The Plant Journal** 4(6), 1051-1061.
- Holden, J., Suzuki, D.T. (1968). Dominant temperature-sensitive (DTS) lethal mutations on chromosome 3 of *Drosophila melanogaster*. **Genetics** 60, 188-189.

- Holden, J., Suzuki, D.T. (1973). Temperature-sensitive mutations in *Drosophila melanogaster*. XII. The genetic and developmental effects of dominant lethals on chromosome 3. **Genetics** 73, 445-458.
- Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, T.S., Young, R.A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. **Cell** 95, 717-728.
- Homyk, T. Jr., Szidonya, J. & Suzuki, D.T. (1980). Behavioral mutants of *Drosophila melanogaster*. III. Isolation and mapping of mutations by direct visual observations of behavioral phenotypes. **Mol. Gen. Genet.** 177, 553-565.
- Hong, S.W., Vierling, E. (2000). Mutants of *Arabidopsis thaliana* defective in the acquisition of tolerance to high temperature stress. **Proc. Natl. Acad. Sci. USA** 97(8), 4392-4397.
- Hopp T.P. and Woods K.R. (1981). Prediction of protein antigenic determinants from amino acid sequences. **Proc. Natl. Acad. Sci. USA** 78 (6), 3824-3828.
- Horowitz, N.H. and Leupold, U. (1951). Some recent studies bearing on the one gene-one enzyme hypothesis. **Cold Spring Harbor Symp. Quant. Biol.** 51, 65-74.
- Horvitz, H.R. and Sulston, J.E. (1980). Isolation and genetic characterization of cell-lineage mutants of the nematode *Caenorhabditis elegans*. **Genetics** 96:435-454.
- Huang, B., Rifkin, M.R., Luck, D.J.L. (1977). Temperature-sensitive mutations affecting flagellar assembly and function in *Chlamydomonas reinhardtii*. **J. Cell Biol.** 72, 67-85.
- Huang, E., Ishida, S., Pittman, J., Dressman, H., Bild, A., Kloos, M., D'Amico, M., Pestell, R., West, M., Nevins, J. (2003). Gene expression phenotypic models that predict the activity of oncogenic pathways. **Nat. Genet.**, 34, 226-230.
- Hubank, M., and Schatz, D. (1994). Identifying difference in mRNA expression by representational difference analysis of cDNA. **Nucleic Acids Res.** 22, 5640-5648.
- Hughes T.R. et al. (2000). Functional Discovery via a compendium of expression profiles. **Cell** 102, 109-126.
- Humphery-Smith, I., Blackstock, W. (1997). Proteome analysis: genomics via the output rather than the input code. **Journal of Protein Chemistry** 16(5), 537-544.
- Hung C.-Y., Lin Y., Zhang M., Pollock S., Marks M.D., and Schiefelbein J. (1998). A common position-dependent mechanism controls cell-type patterning and GLABRA2 regulation in the root and hypocotyl epidermis of *Arabidopsis*. **Plant Physiol.** 117, 73-84.
- Hurlstone, A.F.L., Haramis, A-P., G., Wienholds, E., Begthel, H., Korving, J., van Eeden, F., Cuppen, E., Zivkovic, D., Plasterk, R.H.A., Clevers, H. (2003). The Wnt/ β -catenin pathway regulates cardiac valve formation. **Nature** 425, 633-637.
- Hutchison, C. A., Peterson, S. N., Gill, S. R., Cline, R. T., White, O., Fraser, C. M., Smith, H. O. & Venter, J. C. (1999). Global transposon mutagenesis and a minimal *Mycoplasma* genome. **Science** 286, 2165-2169.
- Ikonomov, O.C., Jacob, M.H. (1996). Differential display protocol with selected primers that preferentially isolates mRNAs of moderate-to-low abundance in a microscopic system. **Biotechniques** 20, 1030-1042.
- Irion, U. & Leptin, M. (1999). Developmental and cell biological functions of the *Drosophila* DEAD-box protein abstrakt. **Curr. Biol.** 9, 1373-1381.
- Ito, T., Kim, G.T., Shinozaki, K. (2000). Disruption of an *Arabidopsis* cytoplasmic ribosomal protein S13-homologous gene by transposon-mediated mutagenesis causes aberrant growth and development. **The Plant Journal** 22(3), 257-264.
- Jack, T., Brockman, L. L. and Meyerowitz, E. M. (1992). The homeotic gene APETALA3 of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. **Cell** 68, 683-697.

- Jack, T., Fox, G. L. and Meyerowitz, E. M. (1994). *Arabidopsis* homeotic gene *APETALA3* ectopic expression: transcriptional and posttranscriptional regulation determine floral organ identity. **Cell** 76, 703-716.
- Jaeken, J., Stibler, H., Hagberg, B. (eds). (1991). The Carbohydrate-Deficient Glycoprotein Syndrome: A new inherited multisystemic disease with severe nervous system involvement. **Acta Paediatr. Scand.** 375 Suppl., 1-71.
- Jaeken, J., Matthijs, G., Carchon, H., Van Schaffingen E. (2001). Defects of N-glycan synthesis. In Scriver CR, Beaudet AL, Sly WS and Valle D (eds.), **The metabolic and molecular bases of inherited disease** 8th ed., McGraw-Hill, New York.
- James, D.W. and Dooner, H.K. (1990). Isolation of EMS-induced mutants in *Arabidopsis* altered in seed fatty acid desaturation. **Theor. Appl. Genet.** 80, 241-245.
- Jander, G., Norris, S.R., Rounsley, S.D., Bush, D.F., Levin, I.M., Last, R.L. (2002). *Arabidopsis* map-based cloning in the post-genome era. **Plant Physiology** 129,440-450.
- Jeffreys, A.J., Wilson, V., Thein, S.L. (1985). Hypervariable 'minisatellite' regions in human DNA. **Nature** 314, 67-73.
- Ji, Y., Zhang, B., Van Horn, S.F., Warren, P., Woodnutt, G., Burnham, M.K.R., and Rosenberg, M. (2001). Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. **Science** 43, 2266-2269.
- Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V., and Kim, S.K. (2001). Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. **Proc. Natl. Acad. Sci. USA** 98, 218-223.
- Jiao, Y., Yang, H., Ma, L., Sun, N., Yu, H., Liu, T., Gao, Y., Gu, H., Chen, Z., Wada, M., et al. (2003). A genome-wide analysis of blue-light regulation of *Arabidopsis* transcription factor gene expression during seedling development. **Plant Physiol.** 133, 1480-1493.
- Jimenez-Sanchez, G., Childs, B., Valle, D. (2001). Human disease genes. **Nature** 409, 853-855.
- Jones, N., Ougham, H., Thomas, H. (1997). Markers and mapping: we are all geneticists now. **New Phytol.** 137,165-177.
- Judson, N. and Mekalanos, J.J. (2000). TnAraOut, a transposon-based approach to identify and characterize essential bacterial genes. **Nature Biotechnology** 18, 740-745.
- K**amath, R.S., Fraser, A.G., Dong, Y., Poulin, G. et al. (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. **Nature** 421, 231-237.
- Kane, M.D., Jatke, T.A., Stumpf, C.R., Lu, J., Thomas, J.D., Madore, S.J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. **Nucleic Acids Research** 28(22), 4552-4557.
- Karp, A., Seberg, O. Buiatti, M. (1996). Molecular techniques in the assessment of botanical diversity. **Ann. Bot.** (London) 78, 143-149.
- Kellis, J.T. Jr., Nyberg, K., Fersht, A.R. (1989). Energetics of complementary side-chain packing in a protein hydrophobic core. **Biochemistry** 28, 4914-4922.
- Kemphues, K.J., Kusch M., Wolf, N. (1988a). Maternal-effect lethal mutations on linkage group II of *Caenorhabditis elegans*. **Genetics** 120:977-986.
- Kemphues, K.J., Priess, J.R., Morton, D.G., Cheng, N.S. (1988b). Identification of genes required for cytoplasmic localization in early *C. elegans* embryos. **Cell** 52, 311-320.
- Kepes, F. and Schekman, R. (1988) The yeast SEC53 gene encodes phosphomannomutase. **The Journal of Biol.Chem.** 263(19), 9155-9161.
- Kim, D.F., Green, R. (1999). Base-Pairing between 23S rRNA and tRNA in the ribosomal A site. **Molecular Cell** 4, 859-864.

- Kim, H., Shesrud, E.C., Haas, B., Cheung, F., Town, C.D., Quackenbush, J. (2003). Gene expression analyses of *Arabidopsis* Chromosome 2 using a genomic DNA amplicon microarray. **Genome Research** 13,327-340.
- Kim, Y. et al. (2004). Structure and function-based characterization of a new phosphoglycolate phosphatase from *Thermoplasma acidophilum*. **The Journal of Biol.Chem.** 279, 517-526.
- King, J., Fane, B., Haase-Pettingell, C., Mitraki, A., Villafane et al. (1989). Identification of amino acid sequences influencing intracellular folding pathways using temperature-sensitive folding mutations, 225-239 in **Protein Folding, Deciphering the Second half of the Genetic Code**, edited by L. M. Gierasch and J. King A.A.A.S., Washington D.C.
- Kirchbaum, J.B., Claeys, I.V., Nasi, S., Molholt, B., Miller, J.H. (1975). Temperature-sensitive RNA polymerase mutants with altered subunit synthesis and degradation. **Proc. Natl. Acad. Sci. USA** 72, 2375.
- Kiyosue, T., Yamaguchi-Shinozaki, K., Shinozaki, K. (1993). Characterization of cDNA for a dehydration-inducible gene that encodes a CLP A, B-like protein in *Arabidopsis thaliana* L. **Biochem. Biophys. Res. Commun.** 196, 1214-1220.
- Kjaergaard, S. (2004). Congenital disorders of glycosylation type Ia and Ib. **Dan. Med. Bull.** 51,350-63
- Kleinhofs, A., Owais, W. M., Nilan, R. A. (1978). **Mutat. Res.**, 55, 165.
- Knowles, J.R. (1980). Enzyme-catalyzed phosphoryl transfer reactions. **Annu. Rev. Biochem.** 49, 877-919.
- Kobayashi, K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K.K., et al. (2003). Essential *Bacillus subtilis* genes. **Proc. Natl. Acad. Sci. USA** 100 (8), 4678-4683.
- Kohiyama, M., Cousin, D., Ryter, A. & Jacob, F. (1966). Thermosensitive mutants of *Escherichia coli* K 12. I. Isolation and rapid characterization. **Ann. Inst. Pasteur (Paris)** 110, 465-486.
- Kohiyama, M. (1968). DNA synthesis in temperature sensitive mutants of *Escherichia coli*. **Cold Spring Harb. Symp. Quant. Biol.** 33, 317-324.
- Koizumi, M., Yamaguchi-Shinozaki, K., Tsuji, H., Shinozaki, K. (1993). Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in *Arabidopsis thaliana*. **Gene** 129,175-182.
- Koncz, C., Nemeth, K., Redei, G.P., and Schell, J. (1992). T-DNA insertional mutagenesis in *Arabidopsis*. **Plant Mol. Biol.** 20, 963-976.
- Konieczny, A. and Ausubel, F.M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. **Plant J.** 4 (2), 403-10.
- Konishi, M., Sugiyama, M. (2003). Genetic analysis of adventitious root formation with a novel series of temperature-sensitive mutants of *Arabidopsis thaliana*. **Development** 130, 5637- 5647.
- Kononov, M.E., Bassuner, B., Gelvin, S.B. (1997). Integration of T-DNA binary vector "backbone" sequences into the tobacco genome: evidence for multiple complex patterns of integration. **Plant J** 11: 945-957.
- Koonin, E.V., A.R. Mushegian, R.L. Tatusov, S.F. Altschul, S.H. Bryant, P. Bork, Valencia, A. (1994). Eukaryotic translation elongation factor 1 gamma contains a glutathione transferase domain—Study of a diverse, ancient protein superfamily using motif search and structural modeling. **Protein Sci.** 3, 2045-2054.
- Koomneef, M., Dellaert, L.M.W., van der Veen, J.H. (1982). EMS and radiation-induced mutation frequencies at individual loci in *Arabidopsis thaliana* (L.) Heynh. **Mutation Research**, 93, 109.
- Koomneef, M., Reuling, G. and Karssen, C.M. (1984). The isolation and characterization of abscisic acid insensitive mutants of *Arabidopsis thaliana*. **Physiol. Plant.** 61, 377-383.
- Koomneef, M., Bentsink, L., Hilhorst, H. (2002). Seed dormancy and germination. *Curr. Opin. Plant Biol.* 5 (1), 33-36.
- Kooter, J.M., and Mol, J.N.M. (1993). Trans-inactivation of gene expression in plants. *Curr. Opin. Biotechnol.* 4,166-171.
- Krasnewich, D., Holt, G.H., Brantly, M., Skovby, F., Redwine, J. and Gahl, W. (1995). Abnormal synthesis of dolichol-linked oligosaccharides in carbohydrate-deficient glycoprotein syndrome. **Glycobiology**, 5, 503-510.

- Krauskopf, A., Blackburn, E.H. and Blackburn, E.H. (1996). Control of telomere growth by interactions of RAP1 with the most distal telomeric repeats. **Nature** 383, 354-357.
- Kreps, J.A., Wu, Y., Chang, H.S., Zhu, T., Wang, X., Harper, J.F. (2002) Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. **Plant Physiol.** 130, 2129-2141.
- Krizek, B. A. and Meyerowitz, E. M. (1996). The *Arabidopsis* homeotic genes *APETALA3* and *PISTILLATA* are sufficient to provide the B class organ identity function. **Development** 122, 11-22.
- Krysan, P.J., Young, J.C., Sussman, M.R. (1999) T-DNA as an insertional mutagen in *Arabidopsis*. **Plant Cell** 11, 2283-2290.
- Kuhn, E. (2001). From library screening to microarray technology: strategies to determine gene expression profiles and to identify differentially regulated genes in plants. **Annals of Botany** 87, 139-155.
- Kurata, N., Nagamura, Y., Yamamoto, K., Harushima, Y., Sue, N., et al. (1994b). A 300 kilobase interval genetic map of rice including 883 expressed sequences. **Nature Genetics** 8, 365-372.
- Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C., Sninsky, J.J., (1990). Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. **Nucleic Acids Res.** 18, 999-1005.
- Lad, C. et al. (2003). The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. **Proc. Natl. Acad. Sci. USA** 100, 5607-5610.
- Lahiri, S.D., Zhang, G., Radstrom, P., Dunaway-Mariano, D., Allen, K.N. (2002). Crystallization and preliminary X-ray diffraction studies of β -phosphoglucosyltransferase from *Lactococcus lactus*. **Acta Cryst.** D58,324-326.
- Lahiri, S.D. et al. (2003). The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. **Science** 299, 2067-2071.
- Lahiri, S.D., Zhang, G., Dunaway-Mariano, D., Allen, K.N. (2004). Analysis of the substrate specificity loop of the HAD superfamily cap domain. **Biochemistry** 43(10), 2812-20.
- Laibach, F. (1943) *Arabidopsis thaliana* (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen. **Bot. Archiv.** 44, 439-455.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., et al. (2001). Initial sequencing and analysis of the human genome. **Nature** 409, 860-921.
- Lane, D.R., Wiedemeier, A., Peng, L., Hofte, H., Vernhettes, S., Desprez, T., Hocart, C.H., Birch, R.J., Baskin, T.I., Burn, J.E., Arioli, T., Betzner, A.S. & Williamson, R.E. (2001). Temperature-sensitive alleles of RSW2 link the KORRIGAN endo-1,4- β -glucanase to cellulose synthesis and cytokinesis in *Arabidopsis*. **Plant Physiol.** 126, 278-288.
- Langdale, J.A., Metzler, M.C. & Nelson T. (1987). The argentinia mutation delays normal development of photosynthetic cell-types in *Zea Mays*. **Dev. Biol.** 122, 243-255.
- Lashkari, D.A., DeRisi, J.L., McCusker, J.H., Namath, A.F., Gentile, C., Hwang, S.Y., Brown, P.O., Davis, R.W. (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. **Proc.Natl.Acad.Sci.USA** 94, 13057-13062.
- Laufs, P., Autran, D., and Traas, J. (1999) A chromosomal paracentric inversion associated with T-DNA integration in *Arabidopsis*. **Plant J.** 18, 131-139.
- Lemieux, B., Miquel, M., Somerville, C. and Browse, J. (1990). Mutants of *Arabidopsis* with alterations in seed lipid fatty acid composition. **Theor. Appl. Genet.** 80, 234-240.
- Lemieux, B., Aharoni, A. & Schena M. (1998). Overview of DNA chip technology. **Molecular Breeding** 4, 277-289.
- Lenhard, M., Jürgens, G., Laux, T. (2002). The *WUSCHEL* and *SHOOTMERISTEMLESS* genes fulfil complementary roles in *Arabidopsis* shoot meristem regulation. **Development** 129, 3195- 3206.

- Leung, J., Bouvier-Durand, M., Morris, P.C., Guerrier, D., Cheddor, F., Giraudat, J. (1994). *Arabidopsis* ABA response gene *ABI1*: features of a calcium-modulated protein phosphatase. **Science** 264, 1448-1452.
- Levy, F., Johnsson, N., Rumenapf, T., Varshavsky, A. (1996). **Proc. Natl. Acad. Sci. USA** 93, 4907-4912.
- Leyser, H.M.O., Furner, I.J. (1992). Characterisation of three shoot apical meristem mutants of *Arabidopsis thaliana*. **Development** 116, 397-403.
- Li, S.L. and Redei, G.P. (1969). Estimation of mutation rates in autogamous diploids. **Rad. Bot.** 9, 125-131.
- Li, J., Last, R.L. (1996). The *Arabidopsis thaliana* *trp5* mutant has a feedback-resistant anthranilate synthase and elevated soluble tryptophan. **Plant Physiol.** 110, 51-59.
- Liang, P. and Pardee, A.B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. **Science** 257 (5072), 967-971.
- Lightner, J., James, D.W. Jr., Dooner, H.K., Browse, J. (1994). Altered body morphology is caused by increased stearate levels in a mutant of *Arabidopsis*. **Plant J.** 6, 401-412.
- Lin, X. et al. (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. **Nature** 402, 761-768.
- Linskens, M.H., Feng, J., Andrews, W.H., Enlow, B.E., Saati, S.M., Tonkin, L.A., Funk, W.D., Villeponteau, B. (1995). Cataloging altered gene expression in young and senescent cells using enhanced differential display. **Nucleic Acids Res.** 23(16), 3244-3251.
- Lipshutz, R.J., Fodor, S.P., Gingeras, T.R., Lockhart, D.J. (1999). High density synthetic oligonucleotide arrays. **Nat. Genet.** 21, 20-24.
- Lister, C. & Dean, C. (1993). Recombinant Inbred Lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. **Plant J.** 4, 745-750.
- Liu, Y.G., Mitsukawa, N., Oosumi, T., Whittier, R.F. (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. **Plant J.** 8, 457-463.
- Liu, Y.G., Shirano, Y., Fukaki, H., Yanai, Y., Tasaka, M., Tabata, S., and Shibata, D. (1999). Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. **Proc. Natl. Acad. Sci. USA** 96, 6535-6540.
- Liu, Y., Schiff, M., Serino, G., Deng, X.-W., and Dinesh-Kumar, S.P. (2002b). Role of SCF ubiquitin-ligase and the COP9 signalosome in the *N* gene-mediated resistance response to Tobacco mosaic virus. **Plant Cell** 14, 1483-1496.
- Livak, K.J. (1999). Allelic discrimination using fluorogenic probes and the 5' nuclease assay. **Genet. Anal.** 14, 143-149.
- Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. **Nature Biotech.** 14, 1675 - 1680.
- Lockhart, D.J. & Winzeler, E.A. (2000). Genomics, gene expression and DNA arrays. **Nature** 405, 827-836.
- Long, D., and Coupland, G. (1998). Transposon tagging with Ac/Ds in *Arabidopsis*. **Methods Mol. Biol.** 82, 315-328.
- Long, J.A., Woody, S., Pothig, S., Meyerowitz, E.M., Barton, K.M. (2002). Transformation of shoots into roots in *Arabidopsis* embryos mutant at the *TOPELESS* locus. **Development** 129, 2297-2306.
- Lo Schiavo, F., Giuliano, G., de Vries, S.C., Genga, A., Bollini, R., Pitto, L., Cozzani, F., Nuti-Ronchi, V., Terzi, M. (1990). A carrot cell variant temperature-sensitive for somatic embryogenesis reveals a defect in the glycosylation of extracellular proteins. **Mol. Gen. Genet.** 223, 385-393.
- Lo Schiavo, F., Giuliano, G., Sung, Z.R. (1988). Characterization of a temperature-sensitive carrot cell mutant impaired in somatic embryogenesis. **Plant Sci.** 54, 157-164.
- Lukowitz, W., Gillmor, C.S. & Scheible, W.R. (2000). Positional cloning in *Arabidopsis*. Why it feels good to have a genome initiative working for you. **Plant Physiol.** 123, 795-805.

- Machida, C., Onouchi, H., Koizumi, S., Hamada, S., Semiarti, E. et al. (1997). Characterization of the transposition pattern of the Ac element in *Arabidopsis thaliana* using endonuclease I-SceI. **Proc. Natl. Acad. Sci. USA** 94, 8675-8680.
- Mahalingam, R. et al. (2003). Characterizing the stress/defense transcriptome of *Arabidopsis*. **Genome Biology** 4,R20.
- Maher, E.P. and Martindale, S.J. (1980). Mutants of *Arabidopsis thaliana* with altered responses to auxins and gravity. **Biochem. Genet.** 18, 1041-1053.
- Malmberg, R.L. (1993). Production and analysis of plant mutants, emphasizing *Arabidopsis thaliana*. In: **Methods in Plant Molecular Biology and Biotechnology**, B.R. Glick and J.E. Thompson, eds (Boca Raton, FL: CRC Press).
- Markwell, J.P., Danko, S.J., Bauwe, H., Osterman, J., Gorz, H.J., Haskins, F.A. (1986). A temperature-sensitive chlorophyll b-deficient mutant of sweetclover (*Melilotus alba*). **Plant Physiol** 81, 329-334.
- Markwell, J., Osterman, J.C. (1992). Occurrence of temperature-sensitive phenotypic plasticity in chlorophyll-deficient mutants of *Arabidopsis thaliana*. **Plant Physiol** 98: 392-394.
- Marra, M. et al. (1999). zA map for sequence analysis of the *Arabidopsis thaliana* genome. **Nature Genetics** 22,265-270.
- Marshall, A., Hodgson, J. (1998). DNA chips: an array of possibilities. **Nature Biotechnology** 16, 27-31.
- Martienssen, R.A. (1998) Functional genomics: probing plant gene function and expression with transposons. **Proc. Natl. Acad. Sci. USA** 95,2021-2026.
- Martin, K.J. and Pardee, A.B. (1999). Principles of differential display. **Methods Enzymol.** 303, 234 -258.
- Martin, J., Horwich, A.L., Hartl, F.U. (1992). Prevention of protein denaturation under heat stress by the chaperonin Hsp60. **Science** 258(5084), 995-998.
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., Penny, D. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. **Proc. Natl. Acad. Sci. USA** 99, 12246-12251.
- Marton, M.J., DeRisi, J.L., Bennett, H.A., Iyer, V.R., Meyer, M.R., Roberts, C.J., Stoughton, R., Burchard, J., Slade, D., Dai, H. (1998). Drug target validation and identification of secondary drug target effects using DNA microarrays. **Nat. Med.** 4, 1293-1301.
- Mathé, C., Déhais, P., Pavy, N., Rombauts, S., Van Montagu, M., Rouzé, P. (2000). Gene prediction and gene classes in *Arabidopsis thaliana*. **Journal of Biotechnology** 78, 293-299.
- Matthijs, G., Schollen, E., Pirard, M., Budarf, M.L., Van Schaftingen, E., Cassiman, J.J. (1997). PMM(PMM1), the human homologue of SEC53 or Yeast phosphomannomutase, is localized on chromosome 22q13. **Genomics** 40, 41-47.
- Matz, M.V. and Lukyanov, S.A. (1998). Different strategies of differential display: areas of application. **Nucleic Acids Res.** 26 (24), 5537-5543.
- Mayer, U., Torres-Ruiz, R.A., Berleth, T., Miséra, S. and Jürgens, G. (1991). Mutations affecting body organisation in the *Arabidopsis* embryo. **Nature** 353, 402-407.
- Mayoh, H., Suzuki, D.T. (1973). Temperature-sensitive mutations in *Drosophila melanogaster* XVI. The genetic properties of sex-linked recessive cold-sensitive mutants. **Can. J. Genet. Cytol.**
- McCallum, C.M., Comai, L., Greene, E.A., Henikoff, S. (2000). Targeted screening for induced mutations. **Nature Biotechnology** 18,455-457
- McClelland, M., Mathieu-Daude, F., and Welsh, J. (1995). RNA fingerprinting and differential display using arbitrarily primed PCR. **Trends Genet.** 11, 242-246.
- McEachern, M.J., Iyer, S., Fulton, T.B., Blackburn, E.H. (2000a). Telomere fusion caused by mutating the terminal region of telomeric DNA. **Proc. Natl. Acad. Sci. USA** 97, 11409-11414.

- McElver, J., Tzafrir, I., Aux, G., Rogers, R. et al. (2001). Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. **Genetics** 159, 1751-1763.
- McGavin, S. (1968). Interallelic complementation and allostery. **J.Mol. Biol.** 37, 239-242.
- McKelvie, A.D. (1962). A list of mutant genes in *Arabidopsis thaliana* (L.) Heynh. **Radiat. Bot.** 1, 233-241.
- Meinke, D.W., Cherry, J.M., Dean, C., Rounsley, S.D., Koornneef, M. (1998). *Arabidopsis thaliana*: a model plant for genome analysis. **Science** 282 662,679-682.
- Meinke, D.W. and Sussex, I.M. (1979). Embryo-lethal mutants of *Arabidopsis thaliana*: a model system for genetic analysis of plant embryo development. **Dev. Biol.** 72, 50-61.
- Meinke, D.W., Meinke, L.K., Showalter, T.C., Schissel, A.M., Mueller, L.A., Tzafrir, I. (2003). A sequence-based map of *Arabidopsis* genes with mutant phenotypes. **Plant Physiol.** 131, 409-418.
- Mekhedov, S., Martinez de Ilarduya, O., Ohlrogge, J. (2000). Towards a functional catalog of the plant genome: a survey of genes for lipid biosynthesis. **Plant Physiol.** 122, 389-401.
- Meyerowitz, E.M. (1987). *Arabidopsis thaliana*. **Ann.Rev.Genet.** 21,93-111.
- Meyerowitz, E.M. (2001). Prehistory and History of Arabidopsis Research. **Plant Physiol.** 125, 15-19.
- Meyerowitz, E.M. (2002). Plants Compared to Animals: The Broadest Comparative Study of Development. **Science** 295,1482-1485.
- Meyer, K., Leube, M.P., Grill, E. (1994). A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. **Science** 264, 1452-1455.
- Meyers, B.C., Scalabrin, S., Morgante, M. (2004). Mapping and sequencing complex genomes: let's get physical! **Nature reviews genetics** 5, 578-588.
- Michaels, S.D. and Amasino, R.M. (1998). A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. **Plant J.** 14, 381-385.
- Michaels, G.S. and Carr, D.B. (1998). Cluster analysis and data visualization of large-scale gene expression data. **Pacific Symposium on Biocomputing** 3, 42-53.
- Michelmore, R.W., Paran, I. and Kesseli, R.V. (1991). Identification of markers linked to disease-resistance genes by bulked segregation analysis: A rapid method to detect markers in specific genome regions by using segregating populations. **Proc. Natl. Acad. Sci. U.S.A.** 88, 9828-9832.
- Mildvan, A.S. (1997). Mechanisms of signaling and related enzymes. **Proteins** 29, 401-416.
- Miller, J.H. (1972). Isolation and mapping of temperature-sensitive lethal mutants. In: Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, New York: 235.
- Mindrinis, M., Katagiri, F., Yu, G.L., Ausubel, F.M. (1994) The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. **Cell** 78, 1089-1099
- Mitchell, H.K. and Houlahan, M.D. (1946). Adenine-requiring mutants of *Neurospora crassa*. **Federation Proc.** 5, 370-375.
- Mitraki, A., Fane, B., Haase-Pettingel, C., Sturtevant, J., King, J. (1991). Global suppression of protein folding defects and inclusion body formation. **Science** 5, 54-58.
- Mitraki, A., Danner, M., King, J., and Seckler, R. (1993). Temperature-sensitive mutations and second-site suppressor substitutions affect folding of the P22 tailspike protein *in vitro*. **J. Biol. Chem.** 268, 20071-20075.
- Miwa, J., Schierenberg, E., Miwa, S., Von Ehrenstein, G. (1980). Genetics and mode of expression of temperature-sensitive mutations arresting embryonic development in *Caenorhabditis elegans*. **Dev. Biol.** 76,160-174.
- Mohan, M., Nair, S., Bhagwat, A., Krishna, T.G., Yano, M., Bhatia, C.R., Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. **Molecular Breeding** 3, 87-103.

- Moore, J.M., Vielle-Calzada, J.-P., Gagliano, W. and Grossniklaus, U. (1997). Genetic characterization of *hadad*, a mutant disrupting female gametogenesis in *Arabidopsis thaliana*. **Cold Spring Harb. Symp. Quant. Biol.** 62, 35-47.
- Mootz, H.D., Muir, T.W. (2002). Protein splicing triggered by a small molecule. **J. Am. Chem. Soc.** 124, 9044-9045.
- Morais, M.C., Zhang, W., Baker, A.S., Zhang, G., Dunaway-Mariano, D., Allen, K.N. (2000). The crystal structure of *Bacillus cereus* phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. **Biochemistry** 39, 10385-10396.
- Morris, E.R. and Walker, J.C. (2003). Receptor-like protein kinases: the keys to response. **Curr. Opin. Plant Biol.** 6 (4):339-342.
- Martin, M.A. & Kaufman, T.C. (1984). Developmental effects of a temperature-sensitive RNA polymerase II mutation in *Drosophila melanogaster*. **Dev. Biol.** 103, 343-354.
- Mozo, T., Dewar, K., Dunn, P., Ecker, J.R., Fischer, S., Kloska, S., Lehrach, H., Marra, M., Martienssen, R., Meier-Ewert, S., Altmann, T. (1999). A complete BAC-based physical map of the *Arabidopsis thaliana* genome. **Nature Genetics** 22, 271-275.
- Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. **Physiol Plant** 18, 100-127.
- Murgola, E.J., Pagel, F.T., Hijazi, K.A., Arkov, A.L., Xu, W. and Zhao, S.Q. (1995). Variety of nonsense suppressor phenotypes associated with mutational changes at conserved sites in *Escherichia coli* ribosomal RNA. **Biochem. Cell Biol.** 73, 925-931.
- Murray, M.G., Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. **Nucleic Acids Res.** 8 (19), 4321-4326.
- Muskett, P.R., Clissold, L., Marocco, A., Springer, P.S., Martienssen, R., Dean, C. (2003). A resource of mapped dissociation launch pads for targeted insertional mutagenesis in the *Arabidopsis* genome. **Plant Physiol.** 132, 506-516.
- Myers et al. (2000). A whole-genome assembly of *Drosophila*. **Science** 287, 2196-2204.

- N**acry, P., Camilleri, C., Courtial, B., Caboche, M., Bouchez, D. (1998). Major chromosomal rearrangements induced by T-DNA transformation in *Arabidopsis*. **Genetics** 149, 641-650.
- Nagy, A., Perrimon, N., Sandmeyer, S., Plasterk, R. (2003). Tailoring the genome: the power of genetic approaches. **Nature Genetics** 33, 276-284.
- Nair, S., Bentur, J.S., Rao, J.R. Mohan, M. (1995). DNA markers tightly linked to a gall midge resistance gene (*Gm2*) are potentially useful for marker-aided selection in rice breeding. **Theor. Appl. Genet.** 91, 68-73.
- Nair, S., Kumar A., Srivastava M.N., Mohan, M. (1996). PCR-based DNA Markers linked to a gall midge resistance gene, *Gm4t*, has potential for marker-aided selection in rice. **Theor. Appl. Genet.** 92, 660-665.
- Nautiyal, S., DeRisi, J.L., Blackburn, E.H. (2002). The genome-wide expression response to telomerase deletion in *Saccharomyces cerevisiae*. **Proc. Natl. Acad. Sci.** 99, 9316-9321.
- Nechiporuk, A., Poss, K.D., Johnson, S.L., Keating, M.T. (2003). Positional cloning of a temperature-sensitive mutant *emmental* reveals a role for *sly1* during cell proliferation in zebrafish fin regeneration. **Dev. Biol.** 258 (2), 291-306.
- Neff, M.M., Chory, J.D., and Pepper, A.E. (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: Experimental applications in *Arabidopsis thaliana* genetics. **Plant J.** 14, 387-392.
- Negishi, T. et al. (2002). cDNA microarray analysis of gene expression during Fe-deficiency stress in barley suggests that polar transport of vesicles is implicated in phytosiderophore secretion in Fe-deficient barley roots. **The Plant Journal** 30(1), 83-94.

- Ng, P.C., Henikoff, S. (2001). Predicting deleterious amino acid substitutions. **Genome Research** 11,863-874
- Ng, P.C., Henikoff, S. (2003). SIFT: predicting amino acid changes that effect protein function. **Nucleic Acids Research** 31(13), 3812-3814.
- Nickle, T.C. and Meinke D.W. (1998). A cytokinesis-deficient mutant of *Arabidopsis* (*cyt1*) characterized by embryonic lethality, incomplete cell walls and excessive callose accumulation. **Plant J.** 15, 321-332.
- Nicol, F., His, I., Jauneau, A., Vernhettes, S., Canut, H., Höfte, H. (1998). A plasma membrane-bound putative endo-1,4- β -D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. **The EMBO Journal** 17(19), 5563-5576.
- Noctor, G., Foyer, C.H. (1998a). Ascorbate and glutathione: keeping active oxygen under control. **Annu. Rev. Plant Physiol. Plant Mol. Biol.** 49, 249-279
- Nordborg, M. et al. (2005). The pattern of polymorphism in *Arabidopsis thaliana*. **PloS Biology** 3(7), e196.
- Noren, C.J., Wang, J. and Perler, F.B. (2000) Angewandte **Chemie Ed. Engl.**, 39
- Nurse, P., Thuriaux, P. and Nasmyth, K. (1976). Genetic control of the cell division cycle in the fission yeast *Schizosaccharomyces pombe*. **Mol. Gen. Genet.** 146, 167-178.
- Nylander, M., Svensson, J., Palva, E.T., Welin, B.V. (2001). Stress-induced accumulation and tissue-specific localization of dehydrins in *Arabidopsis thaliana*. **Plant Mol. Biol.** 45, 263-279.
- O'Connell, K.F., Leys, C.M. & White, J.G. (1998). A genetic screen for temperature-sensitive cell division mutants of *Caenorhabditis elegans*. **Genetics** 149, 1303-1321.
- O'Donovan, G.A. and Ingraham, J.L. (1965). Cold-sensitive mutants of *Escherichia coli* resulting from increased feedback inhibition. **Proc. Natl. Acad. Sci. USA** 54 (2), 451-457.
- Oesterhelt, C., Schnarrenberger, C., Gross, W. (1997). The reaction mechanism of phosphomannomutase in plants. **FEBS Letters** 401,35-37.
- Ohlrogge, J., Benning, C. (2000). Unraveling plant metabolism by EST analysis. **Curr. Opin. Plant Biol.** 3, 224-228.
- Okubo, K. et al. (1992). Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. **Nature Genetics** 2, 173-179.
- Oliver, S.G., Van der Aart, Q.J.M., Agostoni-Carbone M.L., Aigle, M. (1992). The complete DNA sequence of yeast chromosome III. **Nature** 357, 38-46.
- Ornitz, D. M., Moreadith, R. W. and Leder, P. (1991). Binary system for regulating transgene expression in mice: targeting int-2 gene expression with yeast GAL4/UAS control elements. **Proc. Natl. Acad. Sci. USA** 88, 698-702.
- Østergaard, L. and Yanofsky, M.F. (2004). Establishing gene function by mutagenesis in *Arabidopsis thaliana*. **Plant J.** 39, 682-696.
- Ozawa, S., Yasutani, I., Fukuda, H., Komamine, A. & Sugiyama, M. (1998). Organogenic responses in tissue culture of *srd* mutants of *Arabidopsis thaliana*. **Development** 125, 135-142.

- Page, D.R., Grossniklaus, U. (2002). The art and design of genetic screens : *Arabidopsis thaliana*. **Nature reviews genetics** 3, 124-136.
- Pakula, A.A., Young, V.B., Sauer R.T. (1986). Bacteriophage λ cro mutations: effects on activity and intracellular degradation. **Proc. Natl. Acad. Sci. USA** 83 (23), 8829-8833.
- Palladino, M.J., Hadley, T.J., Ganetzky, B. (2002). Temperature-sensitive paralytic mutants are enriched for those causing neurodegeneration in *Drosophila*. **Genetics** 161, 1197-1208.

- Paquette, S.M., Bak, S., Feyereisen, R., (2000). Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome P450 genes of *Arabidopsis thaliana*. **DNA Cell Biol.** 19, 307-317.
- Paran, I. and Michelmore, R.W. (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. **Theor. Appl. Genet.** 85, 985-993.
- Parichy, D.M., Turner, J.M., Parker, N.B. (2003). Essential role for puma in development of postembryonic neural crest-derived cell lineages in zebrafish. **Dev. Biol.** 256 (2), 221-41.
- Parinov, S. and Sundaresan, V. (2000) Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project. **Curr. Opin. Biotechnol.** 11, 157-161.
- Parsell, D.A. and Sauer, R.T. (1989). The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. **J. Biol. Chem.** 264 (13), 7590-7595.
- Pastori, G.M., Kiddle, G., Antoniwi, J., Bernard, S., Veljovic-Jovanovic, S., Verrier, P.J., Noctor, G., Foyer, C.H. (2003). Leaf vitamin C contents modulate plant defense transcripts and regulate genes that control development through hormone signaling. **The Plant Cell** 15, 939-951.
- Paulus, H. (2003). Inteins as targets for potential antimycobacterial drugs. **Front Biosci.** 8, 1157-1165.
- Pavet, V., Olmos, E., Kiddle, G., Mowla, S., Kumar, S., Antoniwi, J., Alvarez, M.E., Foyer, C.H. (2005). Ascorbic acid deficiency activates cell death and disease resistance responses in *Arabidopsis*. **Plant Physiology** 139, 1291-1303.
- Pease, A., Solas, D., Sullivan, E., Cronin, M., Holmes, C., Fodor, S. (1994). Light-directed oligonucleotide arrays for rapid DNA sequence analysis. **Proc. Natl Acad. Sci. USA**, 91, 5022-5026.
- Peixoto, A.A. & Hall, J.C. (1998). Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila*. **Genetics** 148, 827-838.
- Peng, L, Hocart, C.H., Redmond, J.W. & Williamson, R.E. (2000). Fractionation of carbohydrates in *Arabidopsis* root cell walls shows that three radial swelling loci are specifically involved in cellulose production. **Planta** 211, 406-414.
- Pérez-Amador, M.A., Lidder, P., Johnson, M.A., Landgraf, J., Wisman, E., Green, P.J. (2001). New molecular phenotypes in the *dsl* mutants of *Arabidopsis* revealed by DNA microarray analysis. **The Plant Cell** 13, 2703-2717.
- Perler, F.B. (2002). InBase, the intein database. **Nucleic Acids Res.** 28 (1), 344-345.
- Perou, C.M. et al. (2000). Molecular portraits of human breast tumours. **Nature** 406, 747-752.
- Perry, J.A., Wang, T.L., Welham, T.J., Gardner, S., Pike, J.M., Yoshida, S., Parniske, M. (2003). A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. **Plant Physiol.** 131, 866-871.
- Peters, D.G., Kassam, A.B., Yonas, H., O'Hare, E.H., Ferrell, R.E., Brufsky, A.M. (1999). Comprehensive transcript analysis in small quantities of mRNA by SAGELife. **Nucleic Acids Res.** 15, e39.
- Peters, J.L., Constandt, H., Neyt, P., Cnops, G., Zethof, J., Zabeau, M. & Gerats, T. (2001). A physical amplified fragment-length polymorphism map of *Arabidopsis*. **Plant Physiol.** 127, 1579-1589.
- Peters, J.L., Cnudde, F., Gerats, T. (2003). Forward genetics and map-based cloning approaches. **Trends in Plant Science** 8 (10), 484-491.
- Peters, J.L., Cnops, G., Neyt, P., Zethof, J., Cornelis, K., Van Lijsebettens, M., Gerats, T. (2004). An AFLP-based genome-wide mapping strategy. **Theor. Appl. Genet.** 108, 321-327.
- Pickett, F. B., Champagne, M. M. & Meeks-Wagner, D. R. (1996). Temperature-sensitive mutations that arrest *Arabidopsis* shoot development. **Development** 122, 3799-3807.
- Pirard, M., Collet, J.F., Matthijs, G., Van Schaftingen, E. (1997). Comparison of PMM1 with the phosphomannomutases expressed in rat liver and in human cells. **FEBS Letters** 411, 251-254.

- Pirard, M., Achouri, Y., Collet, J.F., Schollen, E., Matthijs, G., Van Schaffingen, E. (1999). Kinetic properties and tissular distribution of mammalian phosphomannomutase isozymes. **Biochem.J.** 339,201-207.
- Ponce, M.R., Robles, P., Micol, J.L. (1999). High-throughput genetic mapping in *Arabidopsis thaliana*. **Mol. Gen. Genet.** 261,408-415.
- Posas, F., Chambers, J.R., Heyman, J.A., Hoeffler, J.P., de Nadal, E., Ariño, J. (2000). The transcriptional response of yeast to saline stress. **The Journ. of Biol.Chem.** 275, 17249-17255.
- Powell, L.D., Paneerselvam, K., Vij, R., Diaz, S., Manzi, A., Buist, N., Freeze, H. and Varki, A. (1994). Carbohydrate-deficient glycoprotein syndrome: not an N-linked oligosaccharide processing defect, but an abnormality in lipid-linked oligosaccharide biosynthesis? **J. Clin. Invest.**, 94, 1901-1909.
- Preston, R.D. (1988) Cellulose microfibril-orienting mechanisms in plant cell walls. **Planta**, 174, 67-74.
- Pries, F., van den Wijngaard, A.J., Bos, R., Pentenga, M., Janssen, D.B. (1994). The role of spontaneous cap domain mutations in Haloalkane dehalogenase specificity and evolution. **The Journal of Biol.Chem.** 269,17490-17494.
- Pruitt, R.E., Meyerowitz, E.M.(1986). Characterization of the Genome of *Arabidopsis thaliana*. **J.Mol.Biol.** 187, 169-183.

- Q**in, L., Overmars, H., Helder, J., Popeijus, H., van der Voort, J., Groenink, W., van Koert, P., Schots, A., Bakker, J., Smant, G. (2000). An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. **Mol. Plant Microbe Interact.** 13, 830-836.
- Quail, P.H. (1998). The phytochrome family: dissection of functional roles and signalling pathways among family members. **Phil. Trans. R. Soc. Lond. B** 353, 1399-1403.
- Quakenbush, J.. (2001). Computational analysis of microarray data. **Nat. Rev. Genet.** 2, 418-427.

- R**ahner, A., Schöler, A., Martens, E., Gollwitzer, B., Schüller, H.-J. (1996) Dual influence of the yeast Cat1p (Snf1p) protein kinase on carbon source-dependent transcriptional activation of gluconeogenic genes by the regulatory gene CAT8. **Nucleic Acids Res** 24, 2331-2337.
- Ramachandran, S. and Sundaresan, V. (2001). Transposons as Tools for Functional Genomics. **Plant Physiology and Biochemistry** 39, 243-252.
- Ramsay, G. (1998). DNA chips: state-of-the-art. **Nature Biotechnology** 16, 40-44.
- Rédei, G.P. (1974). *Arabidopsis thaliana*. In: **Handbook of Genetics**, 151-180; edited by R. C. King. Plenum, New York.
- Rédei, G.P. and Koncz, C. (1992). In: **Methods in Arabidopsis research**, 16; edited by C. Koncz, N-H. Chua, and J. Schell. World Scientific Publishing, Singapore.
- Reeder R.H. (1974). Ribosomes from eukaryotes: genetics. In: **Ribosomes**, M. Nomura, ed. (Cold Spring Harbor, NY), 489-519.
- Regad, F., Lebas, M., Lescure, B. (1994). Interstitial telomeric repeats within the *Arabidopsis thaliana* genome. **J.Mol.Biol.** 239,163-169.
- Renner, C., Trumper, L., Pfitzenmeyer, J.P., Löffin, U., Gerlach, K., Stehle, I., Wadle, A., Pfreundschuh, M. (1998). Differential mRNA display at the single-cell level. **Biotechniques** 24, 720-722.
- Reymond, P., Weber, H., Damond, M. & Farmer, E.E. (2000). Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis*. **Plant Cell** 12, 707-720.
- Rizhsky, L., Liang, H., Mittler, R. (2002). The combined effect of drought stress and heat shock on gene expression in Tobacco. **Plant Physiology** 130,1143-1151.

- Roberts, C.J., Nelson, B., Marton, M.J., Stoughton, R., Meyer, M.R., Bennett, H.A., He, Y.D.D., Dai, H.Y., Walker, W.L., Hughes, T.R., Tyers, M., Boone, C., Friend, S.H. (2000). Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. **Science** 287, 873-880.
- Rose, G. and Roy, S. (1980). Hydrophobic basis of packing in globular proteins. **Proc. Natl. Acad. Sci. USA** 77 (8), 4643-4647.
- Rosenbluth, R., Ezell, D., Suzuki, D.T. (1972). Temperature-sensitive mutations in *Drosophila melanogaster*. IX. Dominant cold-sensitive lethals on the autosomes. **Genetics** 70, 75-88.
- Rounsley, S., Lin, X., Ketchum, K.A. (1998). Large-scale sequencing of plant genomes. **Current Opinion in Plant Biology**, 1, 136-141.
- Rowland, O., Jones, J.D.G. (2001). Unraveling regulatory networks in plant defense using microarrays. **Genome Biology** 2(1), 10011-10013.
- Sablowski, R. W. & Meyerowitz, E. M. (1998). Temperature-sensitive splicing in the floral homeotic mutant *apetala3-1*. **Plant Cell** 10, 1453-1463.
- Sadler, J. R., and Novick, A. (1965). The properties of repressor and the kinetics of its action. **J. Mol. Biol.** 12, 305-327.
- Sagerström, C.G., Sun, B.I., Sive, H.L. (1997). Subtractive cloning: Past, Present, and Future. **Ann. Rev. Biochem.** 66, 751-783.
- Sakurai, H., Mitsuzawa, H., Kimura, M. & Ishihama, A. (1999). The Rpb4 subunit of fission yeast *Schizosaccharomyces pombe* RNA polymerase II is essential for cell viability and similar in structure to the corresponding subunits of higher eukaryotes. **Mol. Cell. Biol.** 19, 7511-7518.
- Salinas-Mondragon, R.E., Garciduenas-Pina, C., and Guzman, P. (1999). Early elicitor induction in members of a novel multigene family coding for highly related RING-H2 proteins in *Arabidopsis thaliana*. **Plant Mol. Biol.** 40, 579-590.
- Sambrook, J., Russell, D.W. (1989). **Molecular Cloning: A Laboratory Manual** 1. Cold Spring Harbor, NY; Cold Spring Harbor Laboratory Press.
- Sanchez-Diaz, A., Kanemaki, M., Marchesi, V., Labib, K. Rapid depletion of budding yeast proteins by fusion to a heat-inducible degenon. **Sci. STKE** 223, 8-9.
- Sargent, T.D. (1987). Isolation of differentially expressed genes. **Methods Enzymol.** 152, 423-432.
- Saurin, A.J., Borden, K.L., Boddy, M.N., and Freemont, P.S. (1996). Does this have a familiar RING? **Trends Biochem. Sci.** 21, 208-214.
- Savelkoul, P.H., Aarts, H.J., de Haas, J., Dijkshoorn, L., Duim, B., Otsen, M., Rademaker, J.L., Schouls, L. & Lenstra J.A.. (1999). Amplified-fragment length polymorphism analysis: the state of an art. **J. Clin. Microbiol.** 37, 3083-3091.
- Schena, M., Shalon, D., Davis, R.W., Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science** 270, 467-470.
- Schena, M. (1996). Genome analysis with gene expression microarrays. **BioEssays** 18, 427-431.
- Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O., Davis, R.W. (1996). Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. **Proc. Natl. Acad. Sci. USA** 93, 10614-10619.
- Schenk, P.M., Kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C., Manners, J.M. (2000). Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. **PNAS** 97(21), 11655-11660.
- Schenk, P.M., Kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C., Manners, J.M. (2000). Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. **Proc. Natl. Acad. Sci. USA** 97, 11655-11660.
- Schlesinger, J. and Levinthal, C. (1963). Hybrid protein formation of *E. coli* alkaline phosphatase leading to *in vitro* complementation. **J. Mol. Biol.** 7: 1-12

- Schmid, K.J., Sørensen, T.R., Stracke, R., Törjék, O., Altmann, T., Mitchell-Olds, T., Weisshaar, B. (2003). Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*. **Genome Research** 13, 1250-1257.
- Schmidt, R., Love, K., West, J., Lenehan, Z., Dean, C. (1997). Description of 31 YAC configs spanning the majority of *Arabidopsis thaliana* chromosome 5. **The Plant Journal** 11(3), 563-572.
- Schmidt, R. (1998). Physical mapping of the *Arabidopsis thaliana* genome. **Plant.Physiol.Biochem.** 36, 1-8.
- Schnable, P.S., Hochholdinger, F., Nakazono, M. (2004). Global expression profiling applied to plant development. **Curr.Opin. in Plant.Biol.** 7, 50-56.
- Schnall, J.A., Cooke, T.J., Cress, D.E. (1988). Genetic analysis of somatic embryogenesis in carrot cell culture: initial characterization of six classes of temperature-sensitive variants. **Dev. Genet.** 9, 49-67.
- Schollen, E., Pardon, E., Heykants, L., Renard, J., Doggett, N.A., Callen, D.F., Cassiman, J.J., Matthijs, G. (1998). Comparative analysis of the phosphomannomutase genes PMM1, PMM2 and PMM2 ψ : the sequence variation in the processed pseudogene is a reflection of the mutations found in the functional gene. **Human Molecular Genetics** 7, 157-164.
- Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lönning, W-E, Saedler, H. & Sommer H. (1992). Characterization of the *Antirrhinum* floral homeotic MADS-box gene *deficiens*: evidence for DNA binding and autoregulation of its persistent expression throughout flower development. **EMBO J.** 11, 251-263.
- Scotti, P.I.D. (1968). A new class of temperature conditional lethal mutants of bacteriophage T4D. **Mutation Res.** 6, 1-14.
- Sedlak B.J. (1986). Spatial and temporal relationships between cell death and tissue overgrowth in imaginal wing disks of the *Drosophila* mutant 1(3)c43hs1. **Dev. Genet.** 6, 199-212.
- Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, Y., Carninci, P., Hayashizaki, Y., Shinozaki, K. (2001). Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. **The Plant Cell** 13, 61-72.
- Seki, M. et al. (2002). Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. **The Plant Journal** 31(3), 279-292.
- Sevastopoulos, C.G., Wehr, C.T., Glaser, D.A. (1977) Large-scale automated isolation of *Escherichia coli* mutants with thermosensitive DNA replication **Proc.Natl.Acad.Sci.USA** 74 (8), 3485-3489.
- Shackelford, G.S., Regni, C.A., Beamer, L.J. (2004). Evolutionary trace analysis of the α -D-phosphohexomutase superfamily. **Protein Science** 13, 2130-2138.
- Sharma, Y.K., Davis, K.R. (1994). Ozone-induced expression of stress-related genes in *Arabidopsis thaliana*. **Plant Physiol.** 105, 1089-1096.
- Shasikant, C.S., Carr, J.L., Bhargava, J., Bentley, K.L., Ruddle, F.H. (1998). Recombinogenic targeting: a new approach to genomic analysis - a review. **Gene** 223,9-20.
- Shen, Z.H., Fox, T.D. (1989). Substitution of an invariant nucleotide at the base of the highly conserved '530-loop' of 15S rRNA causes suppression of yeast mitochondrial ochre mutations. **Nucleic Acids Res.** 17(12), 4535-4539.
- Shenk, T.E., Rhodes, C., Rigby, P.W., Berg, P. (1975). Biochemical method for mapping mutational alterations in DNA with S1 nuclease: the location of deletions and temperature-sensitive mutations in simian virus 40. *Proc. Natl. Acad. Sci. USA* 72 (3), 989-993.
- Shimkets, R.A. et al. (1999). Gene expression analysis by transcript profiling coupled to gene database query. **Nat. Biotechnol.** 17, 798-803.
- Shin, D.H., Roberts, A., Jancarik, J., Yokota, H., Kim, R., Wemmer, D.E., Kim, S.H. (2003). Crystal structure of a phosphatase with a unique substrate binding domain from *Thermotoga maritima*. **Protein Sci.** 12, 1464-1472.

- Shoemaker, D.D., Schadt, E.E., Armour, C.D., He, Y.D., Garrett-Engle, P., McDonagh, P.D., Loerch, P.M., Leonardson, A., Lum, P.Y., Cavet, G. 2001. Experimental annotation of the human genome using microarray technology. **Nature** 409, 922-927.
- Shortle, D., Stites, W.E., Meeker, A.K. (1990). Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. **Biochemistry** 29, 8033-8041.
- Shrader, T. E., Tobias, J. W., and Varshavsky, A. (1993) **J. Bacteriol.** 175, 4364-4374.
- Singer, B.S., Gold, L. (1976) A mutation that confers temperature sensitivity on the translation of rIIb in bacteriophage T4. **J. Mol. Biol.** 103 (3), 627-646.
- Smirnoff, N. (1996). The function and metabolism of ascorbic acid in plants. **Annals of Botany** 78, 661-669.
- Smirnoff, N., Concklin, P., Loewus, F.A. (2001). Biosynthesis of ascorbic acid in plants: a renaissance. **Annu.Rev.Plant.Physiol.Plant.Mol.Biol.** 52,437-467.
- Smith, J. D., Barnett, L., Brenner, S., Russell, R.L. (1970). More mutant tyrosine transfer ribonucleic acids. **J. Mol. Biol.** 54, 1-14.
- Smith, D.H., Berget, P.B., King, J. (1980). Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tailspike protein. I. Fine-structure mapping. **Genetics** 96, 331-352.
- Smith, D.H., King, J. (1981). Temperature-sensitive mutations blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein: inactive polypeptide chains synthesized at 39°C. **J.Mol. Biol.** 145
- Smith, D., Yanai, Y., Liu, Y.G., Ishiguro, S., Okada, K., Shibata, D., Whittier, R.F., Fedoroff, N.V. (1996). Characterization and mapping of Ds-GUS-T-DNA lines for targeted insertional mutagenesis. **Plant J.** 10, 721-732.
- Smith, H.B. (2000). Proteomics. Broad strokes of expressionism? **Plant Cell** 12, 303-304.
- Smits, B.M., Muddek, J., Plasterk, R.H., Cuppen, E. (2004). Target-selected mutagenesis of the rat. **Genomics** 83, 332-334.
- Somerville C. and Somerville S. (1999). Plant functional genomics. **Science** 285, 380-383.
- Sommer, H., Beltran, J.P., Huijser, P., Pape H., Lonig, W.E., Saedler, H., Schwarz-Sommer, Z. (1990). Deficiens, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. **EMBO J.** 9, 605-613.
- Sompayrac, L., Jane, S., Burn, T.C., Tenen, D.G., and Danna, K.J. (1995) Overcoming limitations of the mRNA differential display technique. **Nucleic Acids Res.** 23, 4738-4739.
- Sorlie, T., *et al.* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. **Proc. Natl. Acad. Sci. USA** 98, 10869-10874.
- Sorlie, T., *et al.* (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. **Proc. Natl. Acad. Sci. USA** 100, 8418-8423.
- Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. **J. Mol. Biol.** 98 (3), 503-517.
- Southworth, M.W., Amaya, K., Evans, T.C., Xu, M. and Perler, F.B. (1999) **BioTechniques**, 27, 110-121.
- Spangler, E.A., Blackburn, E.H. (1985). The nucleotide sequence of the 17S ribosomal RNA gene of *Tetrahymena thermophila* and the identification of point mutations resulting in resistance to the antibiotics paromomycin and hygromycin. **J. Biol. Chem.** 260(10), 6334-6340.
- Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. **Mol. Biol. Cell** 9(12), 3273-3297.
- Spiegelman, J.I., Mindrinos, M.N., and Oefner, P.J. (2000) High accuracy DNA sequence variation screening by DHPLC. **BioTechniques** 29, 2-7.

- Springer, P. S., McCombie, W. R., Sundaresan, V. & Martienssen, R. A. (1995). Gene trap tagging of PROLIFERA, an essential MCM2-3-5-like gene in *Arabidopsis*. **Science** 268, 877-880.
- Stears, R.L., Martinsky, T., Schena, M. (2003). Trends in microarray analysis. **Nature Medicine** 9,140-145.
- Steinborn, K., Maulbetsch, C., Priester, B., Trautmann, S., Pacher, T., Geiges, B., Kuttner, F., Lepiniec, L., Stierhof, Y. D., Schwarz, H., Jurgens, G. & Mayer, U. (2002). The *Arabidopsis* PILZ group genes encode tubulin-folding cofactor orthologs required for cell division but not cell growth. **Genes Dev.** 16, 959-971.
- Stemple, D.L. (2004). TILLING - a high-throughput harvest for functional genomics. **Nat. Rev. Genet.** 5, 145-150.
- Stephens, J.L., Brown, S.E., Lapitan, N.L.V., Knudson, D.L. (2004). Physical mapping of barley genes using an ultrasensitive fluorescence in situ hybridization technique. **Genome** 47, 179-189.
- Strachan, T. and Read, A. (1999). **Human molecular genetics** 2. BIOS Scientific New York.
- Stupar, R.M., Lilly, J.W., Town, C.D., Cheng, Z., Kaul, S. *et al.* (2001). Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats. **Proc. Natl. Acad. Sci. USA** 98, 5099-5103.
- Sturtevant, J.M., Yu, M.-H., Haase-Pettingell, C., and King, J. 1989. Thermostability of temperature-sensitive folding mutants of the P22 tailspike protein. **J. Biol. Chem.** 264, 10693-10698.
- Sugihara, J., Baldwin, T.O. (1988). Effects of 3' end deletions from the *Vitrio harvey luxB* gene on luciferase subunit folding and enzyme assembly: generation of temperature-sensitive polypeptide folding mutants. **Biochemistry** 27, 2872-2880.
- Sugimoto, K., Williamson, R.E. and Wasteneys, G.O. (2000). New techniques enable comparative analysis of microtubule orientation, wall texture, and growth rate in intact roots of *Arabidopsis*. **Plant Physiol.** 124, 1493-1506.
- Sugimoto, K., Williamson, R.E., Wasteneys, G.O. (2001). Wall architecture in the cellulose-deficient *rsw1* mutant of *Arabidopsis thaliana*: microfibrils but not microtubules lose their transverse alignment before microfibrils become unrecognizable in the mitotic and elongation zones of roots. **Protoplasma** 215(1-4), 172-83.
- Sugimoto, K., Himmelsbach, R., Williamson, R.E., Wasteneys, G.O. (2003). Mutation or drug-dependent microtubule disruption causes radial swelling without altering parallel cellulose microfibril deposition in *Arabidopsis* root cells. **The Plant Cell** 15, 1414-1429.
- Sugiyama, M. (1999). Organogenesis *in vitro*. **Curr Opin Plant Biol.** 2, 61-64.
- Sugiyama, M. (2003). Isolation and initial characterization of temperature-sensitive mutants of *Arabidopsis thaliana* that are impaired in root redifferentiation. **Plant and Cell Physiology** 44(6), 588-596.
- Sun, P., Enslin, H., Myung, P.S., and Maurer, R.A. (1994). Differential activation of CREB by Ca21-calmodulin-dependent protein kinases type II and type IV involve phosphorylation of a site that negatively regulates activity. **Genes Dev.** 8, 2527-2539.
- Sundaresan, V. (1996). Horizontal spread of transposon mutagenesis: Old elements find new uses. **Trends in Plant Sciences** 1, 184-190.
- Sussman, M.R., Amasino, R.M., Young J.C., Krysan, P.J., Austin-Phillips, S. (2000). The *Arabidopsis* Knockout Facility at the University of Wisconsin-Madison. **Plant Physiol.** 124, 1465-1467.
- Sutcliffe, J.G., Foye, P.E., Erlander, M.G., Hilbush, B.S., Bodzin, L.J., Durham, J.T., Hasel, K.W. (2000). TOGA: an automated parsing technology for analyzing expression of nearly all genes. **Proc. Natl. Acad. Sci. USA** 97, 1976-1981.
- Suzuki, D.T. (1970). Temperature-sensitive mutations in *Drosophila melanogaster*. **Science** 170(959), 695-706.
- Suzuki, D.T., Kaufman, T., Falk, D. and the U.B.C. *Drosophila* Research Group (1976). Conditionally expressed mutations in *Drosophila melanogaster*. In: **The genetics and biology of Drosophila** 1a, 207-263. Edited by E. Novitski and M. Ashburner, Academic Press, New York.

- Suzuki, D.T., Pitternick, I., Hayashi, S., Tarasoff, M., Baillie, D. et al. (1967). Temperature-sensitive mutations in *Drosophila melanogaster*. I. Relative frequencies among gamma-ray and chemically induced sex-linked recessive and semi-lethals. **Proc. Natl. Acad. Sci. USA** 57, 907-912.
- Suzuki, D.T. and Procnier, D. (1969). Temperature-sensitive mutations in *Drosophila melanogaster*. III. Dominant lethals and semilethals on chromosome 2. **Proc. Natl. Acad. Sci. USA** 62, 369-376.
- Swidzinski, J.A, Sweetlove, L.J., Leaver, C.J. (2002). A custom microarray analysis of gene expression during programmed cell death in *Arabidopsis thaliana*. **The Plant Journal** 30(4), 431- 446.
- Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., Stahl, F.W. (1983). The double strand break repair model for recombination. **Cell** 33, 25-35.
- Tai, P.C., Kessler, D.P., Ingraham, J. (1969). Cold-sensitive mutations in *Salmonella typhimurium* which affect ribosome synthesis. **J. Bacteriol.** 97, 1298-1304.
- Tanksley, S.D., Young, D.A., Paterson, H., Bonier Bale, M.W. (1989). RFLP mapping in plant breeding: new tools for an old science. **BioTechnology** 7, 257-264.
- Tanksley, S.D., Ganai, W., Martin, G.B. (1995). Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. **Trends in Genetics** 11(2), 63-68.
- Tantikanjana, T., Mikkelsen, M.D., Hussain, M., Halkier, B.A. and Sundaresan, V. (2004). Functional analysis of the tandem-duplicated P450 genes *SPS/BUS/CYP79F1* and *CYP79F2* in glucosinolate biosynthesis and plant development by *Ds* transposition-generated double mutants. **Plant Physiol.** 135, 840-848.
- Tasaka, E. and Suzuki, D.T. (1973). Temperature-sensitive mutations in *Drosophila melanogaster*. XVII. Heat- and cold-sensitive lethals on chromosome 3. **Genetics** 74, 509-520.
- Terryn, N., Rouze, P. & Van Montagu, M. (1999). Plant genomics. **FEBS. Lett.** 452, 3-6.
- Thangavelu, M., James, A.B., Bankier, A., Bryan, G.J., Dear, P.H., Waugh, R. (2003). HAPPY mapping in a plant genome: reconstruction and analysis of a high-resolution physical map of a 1.9 Mbp region of *Arabidopsis thaliana* chromosome 4. **Plant Biotechnology Journal** 1, 23-31
- The *Arabidopsis* Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. **Nature** 408, 796-815.
- The *C. elegans* Sequencing Consortium (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. **Science** 282, 2012-2018.
- Thimm, O., Essigmann, B., Kloska, S., Altmann, T., Buckhout, T.J. (2001). Response of *Arabidopsis* to iron deficiency stress as revealed by microarray analysis. **Plant Physiology** 127, 1030-1043.
- Thompson, J., Kim, D.F., O'Connor, M., Lieberman, K.R., Bayfield, M.A., Gregory, S.T., Green, R., Noller, H.F., Dahlberg, A.E. (2001). Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. **Proc. Natl. Acad. Sci. USA** 98 (16), 9002-9007.
- Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A. et al. (2003). Large-scale discovery of induced point mutations with high-throughput TILLING. **Genome Research** 13 (3), 524-530.
- Till, B.J., Burtner, C., Comai, L., Henikoff, S. (2004). Mismatch cleavage by single-strand specific nucleases. **Nucleic Acids Res.** 32 (8), 2632-2641.
- Tissier, A., Bourgeois, P. (2001). Reverse genetics in plants. **Current Genomics** 2, 269-284.
- Tobias, J.W., Shrader, T.E., Rocap, G., Varshavsky, A. (1991). The N-end rule in bacteria. **Science** 254, 1374-1377.
- Triman, K., Becker, E., Dammel, C., Katz, J., Mori, H., Douthwaite, S., Yapijakis, C., Yoast, S., Noller, H.F. (1989). Isolation of temperature-sensitive mutant of 16S rRNA in *Escherichia coli*. **J. Mol. Biol.**, 209, 645-653.

- Tröbner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P., Lönig, W.E., Saedler, H., Sommer, H. and Schwarz-Sommer, Z. (1992). *GLOBOSA*: a homeotic gene which interacts with *DEFICIENS* in the control of *Antirrhinum* floral organogenesis. **EMBO J.** 11, 4693-4704.
- Truong, H.T., Pratt, E.A., Rule, G.S., Hsue, P.Y., Ho, C. (1991). Inactive and temperature-sensitive folding mutants generated by tryptophan substitutions in the membrane-bound D-lactate dehydrogenase of *Escherichia coli*. **Biochemistry** 30, 10722-10729.
- Tsai, A.Y., Itoh, M., Streuli, M., Thai, T., Saito, H. (1991). Isolation and characterization of temperature-sensitive and thermostable mutants of the human receptor-like protein tyrosine phosphatase LAR. **J. Biol. Chem.** 266: 10534-10543.
- Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, C., McElver, J., Aux, G., Patton, D., Meinke, D. (2004). Identification of genes required for embryo development in *Arabidopsis*. **Plant Physiol.** 135, 1206-1220.
- Tweedie, S., Bird, A. (2000). Mutant weed breaks silence. **Nature** 405, 137-138.

Ugozzoli, L., Wallace, R.B. (1991). Allele-specific polymerase chain reaction. **Methods Enzymol.** 2, 42-48.

Vaeck, E., Ormenèse, S., Lejeune, P., Adamantidis, A., Jacquard, A., Gerats, A. & Inzé, D. Temperature-sensitive mutants of *Arabidopsis thaliana* impaired in organ development (in prep.)

Valvekens, D., Van Montagu, M., Van Lijsebettens, M. (1988). *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis* root explants using kanamycin selection. **Proc. Natl. Acad. Sci. USA** 87, 5536-5540.

Vandenabeele, S., Van Der Kelen, K., Dat, J., et al. (2003). A comprehensive analysis of hydrogen peroxide-induced gene expression in tobacco. **Proc. Natl. Acad. Sci. USA** 100, 16113-16118.

Van der Veen, J. H. (1966). **Arabidopsis Information Service**, 3, 26.

van Hal, N.L.W. et al. (2000). The Application of DNA microarrays in gene expression analysis. **Journal of Biotechnol.** 78, 271-280.

Van Schaffingen, E. and Jaeken, J. (1995). Phosphomannomutase deficiency is a cause of carbohydrate-deficient glycoprotein syndrome type I. **FEBS Lett.**, 377, 318-320.

van 't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., Schreiber, G.J., Kerkhoven, R.M., Roberts, C., Linsley, P.S., Bernards, R. & Friend S.H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. **Nature** 415, 484-485.

Varadarajan, R., Nagarajaram, H.A., Ramakrishnan, C. (1996). A procedure for the prediction of temperature-sensitive mutants of a globular protein based solely on the amino acid sequence. **Proc.Natl.Acad.Sci.USA** 93, 13908-13913.

Varshavsky, A. (1996a). The N-end rule: functions, mysteries, uses. **Proc.Natl.Acad.Sci.USA** 93, 12142-12149.

Varshavsky, A. (1997). The N-end rule pathway of protein degradation. **Genes to Cells** 2,13-28.

Velculescu, V.E., Zhang, L., Vogelstein, B., Kinzler, K.W. (1995). Serial analysis of gene expression. **Science** 270, 484-487

Veljovic-Jovanovic, S.D., Pignocchi, C., Noctor, G., Foyer, C.H. (2001). Low ascorbic acid in the *vtc-1* mutant of *Arabidopsis* is associated with decreased growth and intracellular redistribution of the antioxidant system. **Plant Physiol.** 127, 426-435.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. **Science** 291, 1304-1351.

Visioli, G., E. Maestri, and N. Marmioli. (1997). Differential display-mediated isolation of a genomic sequence for a putative mitochondrial LMW HSP specifically expressed in condition of induced thermo-tolerance in *Arabidopsis thaliana* (L.) Heynh. **Plant Mol. Biol.** 34, 517–527.

Voet, D. and Voet, J.G. (1995). **Biochemistry**, John Wiley & Sons.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Res.** 23, 4407-4414.

Wada, Y., Yamamoto, M. (1997). Detection of single-nucleotide mutations including substitutions and deletions by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. **Rapid Communications in Mass Spectrometry** 11, 1657-1660.

Walbot, V. (1999). Genes, Genomes, Genomics. What can plant biologists expect from the 1998 National Science Foundation Plant Genome Research Program **Plant Physiol.** 119, 1151-1155.

Wambutt, R. et al. (2000). Progress in *Arabidopsis* genome sequencing and functional genomics. **Journ.of Biotechn.** 78,281-292.

Wang, H., Pang, H., Bartlam, M., Rao, Z. (2005). Crystal structure of human E1 enzyme and its complex with a substrate analog reveals the mechanism of its phosphatase/enolase activity. **J.Mol.Biol.** 348, 917-926.

Wasteneys, G.O. (2000). The cytoskeleton and growth polarity. **Curr. Opin. Plant Biol.** 3, 503–511.

Waterhouse, P.M., M.W. Graham, and M.-B. Wang. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. **Proc. Natl. Acad. Sci. USA** 95, 13959-13964.

Weber, D. and Helentjaris, T. (1989). Mapping RFLP loci in maize using B-A translocations. **Genetics** 121, 583-590.

Wechsler, J. A. and Gross, J. D. (1971). *Escherichia coli* mutants temperature-sensitive for DNA synthesis. **Mol. Gen. Genet.** 113, 273-284.

Weining, S. and Langridge, P. (1991). Identification and mapping of polymorphism in cereal based on the polymerase chain reaction. **Theor. Appl. Genet.**, 82, 209 216.

Weisblum, B. and Davies, J. (1968). Antibiotic inhibitors of the bacterial ribosome. **Bacteriol. Rev.** 32, 493-528.

Welsh, J. and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. **Nucleic Acids Res.** 18, 7213-7218.

Welsh, J., Chada, K., Dalal, S.S., Cheng, R., Ralph, D., McClelland, M., (1992). Arbitrarily primed PCR fingerprinting of RNA. **Nucl. Acids Res.** 20 (19), 4965–4970.

Wen, X., Fuhrman, S., Michaels, G.S., Carr, D.B., Smith, S., Barker, J.L., Somogyi, R. (1998). Large-scale temporal gene expression mapping of central nervous system development. **Proc.Natl.Acad.Sci.USA** 95,334-339.

Wheeler, G.L., Jones, M.A., Smirnov, N. (1998). The biosynthetic pathway of vitamin C in higher plants. *Nature* 393, 365-369.

White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., Martínez de Ilárduya, O., Jaworski, J.G., Ohlrogge, J.B., Benning, C. (2001). A new set of *Arabidopsis* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. **Plant Physiol.** 124, 1582-1594.

Whittington, A.T., Vugrek, O., Wei, K.J., Hasenbein, N.G., Sugimoto, K., Rashbrooke, M.C., Wasteneys, G.O. (2001). MOR1 is essential for organizing cortical microtubules in plants. **Nature** 411, 610-613.

Wiedemeier, A.M., Judy-March, J.E., Hocart, C.H., Wasteneys, G.O., Williamson, R.E. & Baskin T.I. (2002). Mutant alleles of *Arabidopsis* RADially SWOLLEN 4 and 7 reduce growth anisotropy without altering the transverse orientation of cortical microtubules or cellulose microfibrils. **Development** 129, 4821-4830.

Wienholds, E., van Eeden, F., Kusters, M., Mudde, J., Plasterk, R.H., Cuppen, E. (2003). Efficient target-selected mutagenesis in zebrafish. **Genome Res.** 13, 2700–2707.

- Williams, G.K., Kubelikk, A.R., Livaki, J.J., Rafalski, A., Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Res.** 18, 6531-6535.
- Williamson, R.E., Burn, J.E., Birch, R., Baskin, T.I., Arioli, T., Betzner A.S., Cork, A. (2001). Morphology of *rsw1*, a cellulose-deficient mutant of *Arabidopsis thaliana*. **Protoplasma** 215(1-4), 116-27.
- Williamson, R.E., Burn, J.E., Hocart, C.H. (2002). Towards the mechanism of cellulose synthesis. **Trends in Plant Science** 7(10), 461- 467.
- Wisman, E. and Ohlrogge J. (2000). *Arabidopsis* microarray service facilities. **Plant Physiol.** 124, 1468-1471.
- Wolfinger, R.D., Gibson, G., Wolfinger, E.D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C. & Paules, R.S. (2001). Assessing gene significance from cDNA microarray expression data via mixed models. **J. Comput. Biol.** 8, 625-637.
- Wolucka, B.A., Persiau, G., Van Doorselaere, J., Davey, M.W., Demol, H., Vandekerckhove, J., Van Montagu, M., Zabeau, M., Boerjan, W. (2001). Partial purification and identification of GDP-mannose 3', 5'-epimerase of *Arabidopsis thaliana*, a key enzyme of the plant vitamin C pathway. **Proc. Natl. Acad. Sci. USA** 98(26), 14843-14848.
- Wood and Revel (1976). The genome of bacteriophage T4. **Bacteriological Reviews** 40 (4), 847-868.
- Wright, T. R. F. (1968). Phenogenetics of temperature sensitive alleles of lethal myospheroid in *Drosophila*. **Proc. 12th Int. Congr. Genet.** 1, 41.
- X**ie, Z., Kasschau, K.D., and Carrington, J.C. (2003). Negative feedback regulation of dicer-like1 in *Arabidopsis* by microRNA-guided mRNA degradation. **Curr. Biol.** 13, 784-789.
- Y**amamoto, F.-I., Clausen, H., White, T., Marken, J., Hakomoro, S.-I. (1990). Molecular genetic basis of the histo-blood group ABO system. **Nature** 345, 229-233.
- Yamamoto, K., Sasaki, T. (1997). Large-scale EST sequencing in rice. **Plant Molec.Biol** 35, 135-144.
- Yang, C.M., Osterman, J.C., Markwell, J. (1990). Temperature sensitivity as a general phenomenon in a collection of chlorophyll-deficient mutants of sweetclover (*Melilotus alba*). **Biochem.Genet.** (1-2), 31-40.
- Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J. & Speed, T.P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. **Nucleic Acids Res.** 30, e15.
- Yanofsky, M., Ma, H., Bowman, G., Drews, K., Feldmann, K., Meyerowitz, E. (1990). The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcriptional factors. **Nature** 346, 35-39.
- Yassin, A., Fredrick, K., Mankin, A.S. (2005). Deleterious mutations in small subunit ribosomal RNA identify functional sites and potential targets for antibiotics. **Proc.Natl.Acad.Sci.USA** 102(46), 16620-16625.
- Yasutani, I., Ozawa, S., Nishida, T., Sugiyama, M. & Komamine, A. (1994). Isolation of temperature-sensitive mutants of *Arabidopsis thaliana* that are defective in the redifferentiation of shoots. **Plant Physiol.** 105, 815-822.
- Yonezawa, K. and Yamagata H. (1975). **Radiat. Bot.**, 15, 241.
- Yu, G.L. and Blackburn, G.H. (1991). Developmentally programmed healing of chromosomes by telomerase in Tetrahymena. **Cell** 67(4), 823-832.
- Yu, M.-H. and King, J. (1984). Single amino acid substitutions influencing the folding pathway of the phage P22 tailspike endorhamnosidase. **Proc.Natl.Acad.Sci.USA** 88, 6584-6588.
- Yuanxin, Y., Chengcai, A., Li, L., Jiayu, G., Guihong, T., Zhangliang, C. (2003). T-linker-specific ligation PCR (T-linker PCR): an advanced PCR technique for chromosome walking or for isolation of tagged DNA ends. **Nucleic Acids Res.** 31 (12), e68.

Yue, H. et al. (2001). An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. **Nucleic Acids Res.** 29, 8 e41.

Zabin, I. and Villarejo, M.R. (1975). Protein Complementation. **Annu. Rev. Biochem.** 44, 295-313.

Zachgo, S., Silva Ede, A., Motte, P., Trobner, W., Saedler, H. & Schwarz-Sommer Z. (1995). Functional analysis of the *Antirrhinum* floral homeotic *DEFICIENS* gene in vivo and in vitro by using a temperature-sensitive mutant. **Development** 121, 2861-2875.

Zeidler, M.P., Tan, C., Bellaiche, Y., Cherry, S., Häder, S., Gayko, U., Perrimon, N. (2004). Temperature-sensitive control of protein activity by conditionally splicing inteins. **Nature Biotechnology** 22, 871-876.

Zhang, P., Hogan, E.L., Bhat, N.R. (1998). Activation of JNK/SAPK in primary glial cultures. II. Differential activation of kinase isoforms corresponds to their differential expression. **Neurochem. Res.** 23, 219-225.

Zhang, M.Q. (1999). Large-scale gene expression data analysis: a new challenge to computational biologists. **Genome Research** 9,681-688.

Zhang, S., Raina, S., Li, H., Li, J., Dec, E., Ma, H., Huang, H., Fedoroff, N.V. (2003). Resources for targeted insertional and deletional mutagenesis in *Arabidopsis*. **Plant Mol. Biol.** 53, 133-150

Zhang, R., Ou, H.Y., Zhang, C.T. (2004). DEG : a database of essential genes. **Nucleic Acids Res.** 32,D271-D272.

Zhao, C., Jung, U.S., Garrett-Engele, P., Roe, T.Y., Cyert, M.S., Levin, D.E. (1998). Temperature-induced expression of yeast *FKS2* is under the dual control of protein kinase C and calcineurin. **Mol. Cell. Biol.** 18, 1013-1022.

Zhu, T., Wang, X. (2000). Large-scale profiling of the *Arabidopsis* transcriptome. **Plant Physiol.** 124, 1472-1476.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., Gruissem, W. (2004). GENEVESTIGATOR. *Arabidopsis* Microarray Database and analysis toolbox. **Plant Physiol.** 136,2621-2632.